

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 14:21:31 ; Search time 1590.15 Seconds
(without alignments)
4125.095 Million cell updates/sec

Title: US-09-743-690-6
Perfect score: 486
Sequence: 1 atgagcacaagtgtgtctca.....tgcgcacacagaagtgga 486

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_estl:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_gss:*
 - 13: em_gss_hum:*
 - 14: em_gss_inv:*
 - 15: em_gss_pln:*
 - 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	378.2	77.8	538	10	BM489973	BM489973 ppg2n.pk0
2	244	50.2	460	9	AI980025	AI980025 pat.pk001
3	91.2	18.8	452	10	BG600657	BG600657 EST505552
4	91.2	18.8	457	10	BI920182	BI920182 EST540117
5	91.2	18.8	457	10	BE922627	BE922627 EST426396
6	89.6	18.4	452	10	BI432561	BI432561 EST535322
7	89.6	18.4	457	10	BE923650	BE923650 EST427419
8	89.6	18.4	458	10	BE472183	BE472183 EST417036
9	89.6	18.4	467	10	BI435772	BI435772 EST538533
10	88	18.1	330	10	BG600209	BG600209 EST505104
11	88	18.1	474	10	BG594050	BG594050 EST492728
12	88	18.1	475	10	BI178711	BI178711 EST519656
13	88	18.1	490	10	BG600210	BG600210 EST505105
14	87	17.9	455	10	BM406970	BM406970 EST581285
15	86.4	17.8	310	9	AW217469	AW217469 EST296183
16	86.4	17.8	378	10	BI432218	BI432218 EST534979
17	86.4	17.8	392	10	BI930261	BI930261 EST550150

18	86.4	17.8	463	10	BE353474	BE353474 EST535851
19	86.4	17.8	464	9	AI771798	AI771798 EST252898
20	86.4	17.8	469	10	BI431821	BI431821 EST534582
21	86.4	17.8	471	9	AI490242	AI490242 EST248568
22	86.4	17.8	493	9	AI486930	AI486930 EST245252
23	86.4	17.8	501	9	AW622766	AW622766 EST306752
24	86.4	17.8	520	10	BF460134	BF460134 06GE06 Ma
c 25	84.8	17.4	692	9	AW616209	AW616209 EST307248
26	84	17.3	222	9	AW038839	AW038839 EST280795
27	84	17.3	460	10	BI435350	BI435350 EST538111
28	84	17.3	476	9	AW929928	AW929928 EST354198
c 30	83.2	17.1	486	9	AW092750	AW092750 EST285930
31	83.2	17.1	182	9	AJ319952	AJ319952 AJ319952
32	83.2	17.1	462	9	AI772725	AI772725 EST253825
33	83.2	17.1	469	9	AI774403	AI774403 EST255503
34	83.2	17.1	472	9	AI772154	AI772154 EST253254
35	83.2	17.1	475	10	BI433692	BI433692 EST536453
36	83.2	17.1	485	10	BI927178	BI927178 EST547067
37	83.2	17.1	489	9	AW040872	AW040872 EST283736
38	83.2	17.1	490	9	AW039971	AW039971 EST282462
39	83.2	17.1	490	9	AW039975	AW039975 EST282466
40	83.2	17.1	492	9	AW093509	AW093509 EST286689
41	82	16.9	487	10	BF153477	BF153477 027A01 Ma
42	81.6	16.8	203	10	BG628048	BG628048 CC-esf1cL
43	81.6	16.8	260	10	BE463103	BE463103 EST354242
44	81.6	16.8	471	9	AW039752	AW039752 EST282225
c 45	80.6	16.6	469	9	AI778944	AI778944 EST259823

ALIGNMENTS

RESULT 1

BM489973	LOCUS	BM489973	538 bp	mrna	linear	EST 07-FEB-2002
DEFINITION	ppg2n.pk002.a9 Normalized Chicken Pituitary/Hypothalamus/Pineal Library (ppg2n) Gallus gallus cDNA clone ppg2n.pk002.a9 5' similar to emb CAC34569.1 (AJ311647) avidin [Gallus gallus], mRNA sequence.					
ACCESSION	BM489973					
VERSION	EST					
KEYWORDS	EST					
SOURCE	chicken.					
ORGANISM	Gallus gallus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.					
AUTHORS	Porter, T.E. and Cogburn, L.A.					
TITLE	ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA Library, USDA/IFAFS Animal Genome Project					
JOURNAL	Unpublished (2002)					
COMMENT	Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2822 Email: cogburnudel.edu, www.Chickest.udel.edu.					
FEATURES	Location/Qualifiers					
source	1. 538					
	/organism="Gallus gallus"					
	/strain="Commercial broiler chickens"					
	/db_xref="taxon:9031"					
	/clone="ppg2n.pk002.a9"					
	/clone.lib="Normalized Chicken Pituitary/Hypothalamus/Pineal Library (ppg2n)"					
	/sex="Male and female"					
	/tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"					
	/dev_stage="Embryonic (d12,d14,d19); post-hatch (1,3,5,7,9 weeks)"					
	/lab_host="E. coli EMDH10B"					
	/note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue at different					

BASE COUNT	140 a	157 c	137 g	104 t
	ages. Single pass sequencing from 5'-end"			

Query Match	77.8%;	Score 378.2;	DB 10;	Length 538;
Best Local Similarity	96.7%;	Prod. No. 9.1e-97;		
Matches 386;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;
QY	88	ceagagatccctgcagaaagtgcctgcctgactgggaaatgacacagatctgggctcc	147	
DB	57	CCCGCCCTCTCTGCCACAAAGTGTCTGCTGACTGGGAAATCGGACACACCTGGGCTCC	116	
QY	148	aacatgaccatcggggctgtgaacagcagaggttgaattcacaggacacatacatcacagcc	207	
DB	117	AACATGACCATCGGGGCTGTGAACAGGAAGGTGAATTACAGGACCTACACACAGCC	176	
QY	208	gtaacagccacatcaaatgagatcaagagctcacattgcattgaggacacaaacacacc	267	
DB	177	GTACAGCCACATCAAAATGAGATCAAGAGGTACACACTGCTATGGAGACACAAACACCATC	236	
QY	268	aacaaggagcccgccacccttggcttcaccgtcaattggaagtgttcagagttccaccc	327	
DB	237	AACAAGAGGCCACAGCCACCTTTGGCTTCACTCTCAATTGGAAGTTTCAGAGTCCACC	296	
QY	328	actgtcttccagggccagtgcttcatagacagggaaatgggaaggaggtcctgaagaacatg	387	
DB	297	ACTGTCTTTCAGGGCCAGTGGCTTCATACAGAGGAACGGGAAGGAGTGCTCTGAAGACCATG	356	
QY	388	tggctgctgcgggtcaagtgttaatgacattggtgatgactggaagctaccagaggtcggc	447	
DB	357	TGGCTGCTGGGGTCAAGTGTTAATGACATTGGTGATGACTGGAAGCTACACAGGGTGGCC	416	
QY	448	atacaatcttcactgcctgcgcacacagaaggagtga	486	
DB	417	ATCAACATCTTCACTCGCTCGGCACACAGAAGGAGTCA	455	

```

RESULT 2
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
SOURCE

AI980025 460 bp mRNA linear EST 07-MAY-2001
pat.pk0010.e3 chicken activated T cell cDNA Gallus gallus cdNA
clone.pat.pk0010.e3 5' similar to avidin, mRNA sequence.

AI980025
AI980025.1 GI:5883053
EST.
chicken.

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 460)
tirunagaru,V.G., Sofer,L., Cui,J. and Burnside,J
An expressed sequence tag database of T-cell-enriched activated
chicken splenocytes: sequence analysis of 5251 clones
Genomics 66 (2), 144-151 (2000)
20318616
Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302 831-1345
Fax: 302-831-3411
Email: joan@udel.edu, www.chickest.udel.edu
Seq primer: r7.
Location/Qualifiers
1..460
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pat.pk0010.e3"
/clone_lib="chicken activated T cell cDNA"
/sex="male"
/cell_type="Con A-activated splenic T cell"

```

```

/lab_host="E.coli_top10_F"
/note="vector: pCDNA3"
BASE COUNT      119 a      118 c      111 g      89 t      23 others
ORIGIN

Query Match      50.28; Score 244; DB 9; Length 460;
Best Local Similarity 83.6%; Pred. No. 1.3e-56;
Matches 280; Conservative 0; Mismatches 50; Indels 5; Gaps 1;

QY      88 ccagagatccctgcacgaaagctgcgtgactcgtggaaatggaccaucgatctgggtccc 147
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      84 CCGCGCTCTCTGCCAGAAAGTCTGCGTGACTGGGAAATGGACCAAGACTGGGTCC 143

QY      148 aacatgaccatcggggctgtgaacagacagaggtgaattcacaggcacacatcacagcc 207
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      144 AACATGACCATCGGGGCTGTGAACGACAGAGGTGAATTACAGAGCAGCTAGATCACAGCC 203

QY      208 gtaacagccacatcaaatgagatcaaaagatcacattgcattgagacacaaacacatc 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      204 GTAACAGCCACATCAATGAGATCAAGAGTCAACCCTGCATGGGANACAAACACCATC 263

QY      268 acaagaggaccgcgccacetttyggttcacgcgtcaattggaagtttccagagtcacacc 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      264 AACAGAGGGCCAGCCACCCTTTGGNTTCACCGNCATATTNGAANTTTTCAGATGCCACC 323

QY      328 actgctctcagggcccagtgctctcatcacagcaggaatgggaaggaggtccctgaagaccatg 387
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      324 ACTGTCMTTACCGGCGCAATCGCTNTATATANAGCAATTGGAGGGGNNTTAGACCA----- 378

QY      388 tggctgctgcggtccaagtgtaatgacattggtga 422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      379 TTNGTNTCTCGGGGCAAAATGTTAANNNTATGGGTGA 413

```

```

RESULT      3
BG600657
LOCUS
DEFINITION BG600657 csts solanum tuberosum cdna clone CSTS29N13 5' sequence,
            mRNA linear EST 12-APR-2001
            452 bp
mRNA sequence.
EST
BG600657.1 GI:13617793
potato.
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 452)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
Bougr,O., Buehl,C.R., Renning,G., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Renning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
Location/Qualifiers
        1..452
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone_lib="CSTS"
        /clone="CSTS29N13"
        /tissue_type="sprouting eyes from tubers"
        /dev_stage="12-14 weeks post harvest"
        /lab_host="SOLR"
        /note="Vector: p Bluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from

```

		/lab_host="E.coli top10 F"					
		/note="vector: pCDNA3"					
BASE COUNT	119 a	118 c	111 g	89 t	23 others		
ORIGIN							
<hr/>							
Query Match	50.2%;	Score 244;	DB 9;	Length 460;			
Best Local Similarity	89.6%;	Pred. No. 1.3e-58;					
Matches 280;	Conservative 0;	Mismatches 50;	Indels 5;	Caps 1;			
<hr/>							
Qy	88	ccagagatcctgcgaagaagtgtcgctgactgctgggaataaggacacacgatctgggtccc	147				
Dd	84	CCCCGCCCTCTGCGCAAAAGTGTTCGGTGACTGGGAATGGACACAGATCTGGGGTCCC	143				
Qy	148	aacatgaccatcggggctgtgaacagcagagagtgtaattcacaggcacctacatcacagcc	207				
Dd	144	AACATGACCATCGGGCGCTGTGAACAGCAGAGAGTGNATTACAGGCACCTACATCACAGCC	203				
Qy	208	gtaacagccaatcaaatgagatcaaaagagtcacacattgcatggagacaaaaaacacatc	267				
Dd	204	GTAACAGCCACATCAAATGAGATCAAAGAGTCAACCATCTGCATGGGANACAACACCATC	263				
Qy	268	acaagaggaccagccacaccttggcttcacgcctaattgaagattttcagagttccacc	327				
Dd	264	AACAAGAGGGCCCCAGCCCCACCTTTGGNTTCACCGCAATTTNGAATTTTCAGATGCCACC	323				
Qy	328	actgtcttcacgggccagtgcttcacagacagaggaatgggaaggagglcctgaadacatg	387				
Dd	324	ACTGCTTTACCGCGCAATTCGTTTATATANAGCAATTTGGAAGGGGNNTTAAGACCA----	378				
Qy	388	tggctgtcgtcggtccaagtgttaatatgacattggtga	422				
Dd	379	TTNCTGTCGGGCAAAATGTTAANNATTGGGTGA	413				
<hr/>							
RESULT	3						
BG600657							
LOCUS							
DEFINITION	EST50552 csts solanum tuberosum cdna clone CSTS29N13 5' sequence,						
ACCESSION	mRNA sequence.						
VERSION	BG600657						
KEYWORDS	EST.						
SOURCE	potato.						
ORGANISM	Solanum tuberosum						
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.						
AUTHORS	van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A., Bougri,O., Buell,C.R., Renning,C., Tanksley,S. and Baker,B.						
TITLE	Generations of ESTs from sprouting potato eyes						
JOURNAL	Unpublished (2000)						
COMMENT	Contact: Cathy Renning The Institute for Genomic Research For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com Seq primer: M13F-R.						
FEATURES	Location/Qualifiers						
source	1..452						
	/organism="Solanum tuberosum"						
	/cultivar="Kennebec"						
	/db_xref="taxon:4113"						
	/clone="CSTS29N13"						
	/clone_lib="CSTS"						
	/tissue_type="sprouting eyes from tubers"						
	/dev_stage="12-14 weeks post harvest"						
	/lab_host="SOLR"						
	/note="Vector: pluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from						

	BASE COUNT	ORIGIN
1	0.0000	0.0000
2	0.0000	0.0000
3	0.0000	0.0000
4	0.0000	0.0000
5	0.0000	0.0000
6	0.0000	0.0000
7	0.0000	0.0000
8	0.0000	0.0000
9	0.0000	0.0000
10	0.0000	0.0000
11	0.0000	0.0000
12	0.0000	0.0000
13	0.0000	0.0000
14	0.0000	0.0000
15	0.0000	0.0000
16	0.0000	0.0000
17	0.0000	0.0000
18	0.0000	0.0000
19	0.0000	0.0000
20	0.0000	0.0000
21	0.0000	0.0000
22	0.0000	0.0000
23	0.0000	0.0000
24	0.0000	0.0000
25	0.0000	0.0000
26	0.0000	0.0000
27	0.0000	0.0000
28	0.0000	0.0000
29	0.0000	0.0000
30	0.0000	0.0000
31	0.0000	0.0000
32	0.0000	0.0000
33	0.0000	0.0000
34	0.0000	0.0000
35	0.0000	0.0000
36	0.0000	0.0000
37	0.0000	0.0000
38	0.0000	0.0000
39	0.0000	0.0000
40	0.0000	0.0000
41	0.0000	0.0000
42	0.0000	0.0000
43	0.0000	0.0000
44	0.0000	0.0000
45	0.0000	0.0000
46	0.0000	0.0000
47	0.0000	0.0000
48	0.0000	0.0000
49	0.0000	0.0000
50	0.0000	0.0000
51	0.0000	0.0000
52	0.0000	0.0000
53	0.0000	0.0000
54	0.0000	0.0000
55	0.0000	0.0000
56	0.0000	0.0000
57	0.0000	0.0000
58	0.0000	0.0000
59	0.0000	0.0000
60	0.0000	0.0000
61	0.0000	0.0000
62	0.0000	0.0000
63	0.0000	0.0000
64	0.0000	0.0000
65	0.0000	0.0000
66	0.0000	0.0000
67	0.0000	0.0000
68	0.0000	0.0000
69	0.0000	0.0000
70	0.0000	0.0000
71	0.0000	0.0000
72	0.0000	0.0000
73	0.0000	0.0000
74	0.0000	0.0000
75	0.0000	0.0000
76	0.0000	0.0000
77	0.0000	0.0000
78	0.0000	0.0000
79	0.0000	0.0000
80	0.0000	0.0000
81	0.0000	0.0000
82	0.0000	0.0000
83	0.0000	0.0000
84	0.0000	0.0000
85	0.0000	0.0000
86	0.0000	0.0000
87	0.0000	0.0000
88	0.0000	0.0000
89	0.0000	0.0000
90	0.0000	0.0000
91	0.0000	0.0000
92	0.0000	0.0000
93	0.0000	0.0000
94	0.0000	0.0000
95	0.0000	0.0000
96	0.0000	0.0000
97	0.0000	0.0000
98	0.0000	0.0000
99	0.0000	0.0000
100	0.0000	0.0000

infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT	159 a	67 c	90 g	151 t
ORIGIN				
Query Match	18.4%;	Score 89.6;	DB 10;	Length 467;
Best Local Similarity	95.8%;	Pred. No. 1e-14;	Mismatches 0;	Gaps 0;
Matches	92;	Conservative	Indels 4;	

QY 1 atggagtcacaaagtgtgctcacatcattgtttcttctttcttgcaactccctttgaaact 60
| | | | |
Db 7 ACGAGCTAAAGTTTGCTCACATCATGTGTTTCTTCTTCCTGCAACTCCTTTGAAACT 66
| | | | |

QY 61 ctcttgaccacgaagaaggatgatgccacagagatc 96
| | | | |
Db 67 CTCTTGGCAGCAAAGAAGAGTAGTGACCAAGATC 102
| | | | |

RESULT 10

BG600209	EST505104	cSTS Solanum tuberosum CDNA clone cSTS2818 5' sequence,	mRNA	linear	EST 12-APR-2001
LOCUS	mRNA sequence.				
DEFINITION	BG600209				
ACCESSION	BG600209.1	GI:13617345			
VERSION	EST.				
KEYWORDS	potato.				
SOURCE	Solanum tuberosum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				

REFERENCE 1 (bases 1 to 330)
Van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)

AUTHORS Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com

TITLE Generations of ESTs from sprouting potato eyes

JOURNAL Unpublished (2000)

COMMENT

FEATURES Source
Location/Qualifiers
1..330
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS2818"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"

NOTES Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

BASE COUNT	107 a	56 c	63 g	104 t
ORIGIN				
Query Match	18.1%;	Score 88;	DB 10;	Length 330;
Best Local Similarity	94.8%;	Pred. No. 2.7e-14;	Mismatches 0;	Gaps 0;
Matches	91;	Conservative	Indels 5;	

QY 1 atggagtcacaaagtgtgctcacatcattgtttcttctttcttgcaactccctttgaaact 60
| | | | |
Db 35 ATGGAGCTAAAGTTTGCTCACATCATGTGTTTCTTCTTCCTGCAACTCCTTTGAAACT 94
| | | | |

QY 61 ctcttgaccacgaagaaggatgatgccacagagatc 96
| | | | |

COMMENT

Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source

Location/Qualifiers
1. 490
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS28110"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

BASE COUNT 160 a 72 c 93 g 165 t
ORIGIN

Query Match 18.1%; Score 88; DB 10; Length 490;

Best Local Similarity 94.8%; Pred. No. 3e-14; Mismatches 5; Indels 0; Gaps 0;

Qy 1 atggagtcacaaagtgtgctcacatcattgtttctttctttctgtcgaactccctttgaaact 60
Db 35 atggagtcacaaagtgtgctcacatcattgtttctttctttctgtcgaactccctttgaaact 94

Qy 61 ctcttgccacgaaagtgatgaccagagatc 96
Db 95 ctcttgccacgaaagtgatgaccagagatc 130

RESULT 14

BM406970

LOCUS BM406970 455 bp mRNA linear EST 22-JAN-2002
DEFINITION EST581285 potato roots Solanum tuberosum cDNA clone cPRO29K4 5' end
mRNA sequence.

ACCESSION BM406970

VERSION BM406970.1 GI:18258588

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 455)

AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,
Utterback,T., Chiemung,A., Bougri,O., Buell,C.R., Ronning,C.,
Tanksley,S. and Baker,B.

Generation of ESTs from potato roots

Unpublished (2001)

CONTACT: Research Genetics, Libraries Division

TEL: 1-800-711-6195

Email: cdna@resgen.com

For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: T3.

FEATURES

source

1. 455
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPRO29K4"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

BASE COUNT 146 a 66 c 89 g 154 t
ORIGIN

Query Match 17.9%; Score 87; DB 10; Length 455;

Best Local Similarity 94.7%; Pred. No. 5.6e-14;
Matches 90; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 tggagtcacaaagtgtgctcacatcattgtttctttctttctgtcgaactccctttgaaactc 61
Db 1 TGGAGTCAAAAGTTTGCACATCATTTGTTTCTTCTTGGCCACTTCCCTTTGAAACTC 60

Qy 62 tcttgccacgaaagtgatgaccagagatc 96

Db 61 TCTTGGCAGCAAAAGATGTGATGGACCAAGATC 95

RESULT 15

AW217469

LOCUS AW217469 310 bp mRNA linear EST 18-MAY-2001
DEFINITION EST296183 tomato flower buds 3-8 mm, Cornell University
Lycopersicon esculentum cDNA clone cTOB1A15 similar to tomato leaf
wound-induced proteinase inhibitor I, mRNA sequence.

ACCESSION AW217469

VERSION AW217469.1 GI:6528343

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 310)

AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
,C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

Generation of ESTs from tomato flower tissue

Unpublished (1999)

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1. 310
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB1A15"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/tissue_type="flower"
/dev_stage="3-8mm buds"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

BASE COUNT 103 a 57 c 54 g 96 t
ORIGIN

Query Match 17.8%; Score 86.4; DB 9; Length 310;

Best Local Similarity 93.8%; Pred. No. 7.5e-14;

Matches 90; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 atggagtcacaaagtgtgctcacatcattgtttctttctttctgtcgaactccctttgaaact 60

db 25 ATGGAGTCAAAAGTTTGGTCACATCATCTTTCTTCTTGTGCACTTCCTTTGAACT 84
Oy 61 ntottggcagcaaaagaaagtgtatggaccagagatc 96
DB 85 CTCATGGCAGCAAAAGAAATTGATGGGCCAGAACTC 120

Search completed: September 20, 2002, 15:50:02
Job time: 5311 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	389.4	80.1	604	4	US-08-831-399-3	Sequence 3, Appli	
2	279.8	57.6	484	1	US-08-554-586-1	Sequence 1, Appli	
C	3	37	7218	1	US-08-232-463-14	Sequence 14, Appl	
	4	31.4	6.5	8040	1	US-08-596-291-1	Sequence 1, Appli
	5	31.4	6.5	8040	3	US-09-100-804-1	Sequence 1, Appli
6	31.4	6.5	8043	5	PCR-US94-09943-1	Sequence 1, Appli	
7	31	6.4	1662	2	US-08-956-012-2	Sequence 2, Appli	
C	8	30.8	6.3	954	1	US-08-170-294-1	Sequence 1, Appli
	9	30.8	6.3	954	2	US-08-664-855-1	Sequence 1, Appli
	10	30.8	6.3	954	3	US-09-049-289-1	Sequence 1, Appli
11	30.6	6.3	400	4	US-09-022-238-1	Sequence 1, Appli	
12	30.6	6.3	1644	1	US-07-675-211-1	Sequence 1, Appli	
13	30.6	6.3	1644	1	US-07-903-047-1	Sequence 1, Appli	
14	30.6	6.3	1644	1	US-08-076-042-1	Sequence 1, Appli	
15	30.6	6.3	1644	2	US-08-757-046A-3	Sequence 3, Appli	
16	30.6	6.3	1644	3	US-09-447-208-3	Sequence 3, Appli	
17	30.6	6.3	1644	3	US-09-135-988-3	Sequence 3, Appli	
18	30.6	6.3	1644	4	US-09-277-716-3	Sequence 3, Appli	
19	30.6	6.3	1644	4	US-08-597-271A-3	Sequence 3, Appli	
20	30.6	6.3	1644	4	US-09-380-061B-13	Sequence 13, Appl	
21	30.6	6.3	1656	3	US-09-111-752-6	Sequence 6, Appli	
22	30.6	6.3	1656	3	US-09-111-752-8	Sequence 8, Appli	
23	30.6	6.3	1656	3	US-09-111-752-9	Sequence 9, Appli	
C	24	30.6	6.3	1779	4	US-09-371-696-1	Sequence 1, Appli
	25	30.6	6.3	3410	4	US-09-020-956-110	Sequence 110, App
	26	30.6	6.3	3410	4	US-09-030-607-110	Sequence 110, App
27	30.6	6.3	3410	4	US-09-439-313-110	Sequence 110, App	

US-08-831-399-3

Query Match	80.1%	Score 389.4	DB 4	Length 604
Best Local Similarity	98.5%	Pred. No. 4.6e-118		
Matches 393	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	88	ccagagatccctgccagaaagtgcctgactgggaatggacacacgactctgggtccc	147	
DB	104	CCCGCCCTCTCTGCCAGAACTGCTGCTGACTGGGAAATGACCAACGATCTGGGCTCC	163	
QY	148	aacatgaccatcgtggggctgtgaacagcagaggttgaattcacaggcaccatcacacagcc	207	
DB	164	AACATGACCATCGGGGGTGTGAACAGACAGAGCTGAATTACAGGACACCTACATCACAGCC	223	
QY	208	gtaacagccacatcaaatgagatcaaaagatccacattgcattggacacacaaaaccatc	267	
DB	224	GTAAAGCCACATCAAAATGAGATCAAAAGATCACCATGTGATGGGACACAAAACCCATC	283	
QY	268	aacagaggacccagccaccctttggcttcaccgtcaattggaagtttccagagtcacacc	327	
DB	284	AACAAGAGGACCCAGCCACCCTTTGGCTTTCACCTCAATTGGAAGTTTTCAGATGCCACC	343	
QY	328	actgctctcagggccagtgctctcatagacaggaagtgggaaggaggtcctctaaagaccatg	387	
DB	344	ACTGCTCTTACGGGCCCACTGGTTCATTACACAGGAAATGGGAAAGGAGCTCTCTGAACACCATG	403	
QY	388	tggtcgtcgtgcagagtggttaatgacatttggttdatgactgggaaggtctaccagaggtcggc	447	
DB	404	TGGCTGCTCGGCTCAAGTGAATGACATTGGTGATGACTGGAAGGCTACACAGGTCGGC	463	
QY	448	ataaacattcttcactgcactgcacacagagaagagtga	486	
DB	464	ATCAACATCTTCATCTGGCCTCGGCACACAGAAAGGAGTGA	502	

RESULT

US-08-554-586-1
 ; Sequence 1, Application US/08554586
 ; Patent No. 5767379
 ; GENERAL INFORMATION:
 ; APPLICANT: BASZCZYNSKI, Chris
 ; APPLICANT: HOOD, Elizabeth
 ; APPLICANT: MADDOCK, Sheila
 ; APPLICANT: MEYER, Terry EuClaire
 ; APPLICANT: REGISTER III, James C.
 ; APPLICANT: WITCHER, Derrick
 ; APPLICANT: HOWARD, John A.
 ; TITLE OF INVENTION: COMMERCIAL PRODUCTION OF AVIDIN IN
 ; TITLE OF INVENTION: WITCHER, Derrick
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/554,586
 ; FILING DATE: 06-NOV-1995
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 33229/352/PIHI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-554-586-1

```

Query Match          57.6%; Score 279.8; DB 1; Length 484;
Best Local Similarity 82.7%; Pred. NO. 3.4e-02;
Matches 320; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 100 qcagaaagtgcctgcctgactgggaataatgaccaaacatctgggctccaacatgaccatc 159
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 GCCAGGAAGTGCTCCCTCACCGGCAAGTGAGCAACATGACCTCGGCTCCAACATGACCATC 143

QY 160 ggggctgtgaacgaagaagtgtaattcacaggaccctacatcacagcgcgttaacagccaca 219
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 GCGCGCTGAACATCCAGGGGCGAGTTCAACCGGCACCTACATCACCGCGGTGACCGCCACC 203

QY 220 tcaaatgagtcaaaagtgatccaccattgcattgggacacaaacacatcaaacagagagacc 279
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 TCCAACAGAGATCAAGGAGTCCCCCTCCACGGTACCCAGAACACCATCAACAGAGGACC 263

QY 280 cagccacccttggctccaccgtcaattgggaagtcttcagagtcacacactgtcttcaag 339
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CAGCGCACCTTCGGGTTCACCGTCAATGTGAAGTTTCCTCAGTCCACCAACCTGTTTCACC 323

QY 340 gggcagtgcttcatagacaggaatgggaagagagtgctctgaagaccatgtgctgctcgg 399
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 GCGCAGTGCTTCATCGAGCCGCAACGGCAAGAGGTGCTCAAGACCATGTGGCTCTCAGG 383

QY 400 tcaagtgttaatgacattgattgatctgaaagctaccagggtggcgcatcaaacattctc 459
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 AGCTCCGTGAATGATCATCGGCGAGGACTGGAAGGCCACCCGCGGTGGGCATCAACATCTTC 443

QY 460 actgcctgcgcacacagaaggagtgca 486
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 444 ACCCGCTCCGCCACCCAGAGGAGCTGA 470

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Iardner
; STREET: 1800 Diagonal Road, Suite 500
; City: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:

```

APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 7.6%; Score 37; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.055;
Matches 16; Conservative 203; Mismatches 168; Indels 0; Gaps 0;

Qy 66 ggcacgaaagaagtgtgacagagatccctgcagaaagtgtcgtgactggaa 125
Db 1426 RR 1367
Qy 126 atggaccaacgtctgggtcccaacatgaccatcggtggtgaaacagagagtgatt 185
Db 1366 RR 1307
Qy 186 cacaggcaccatcacagcgtgaacagccacatcaaatgagatcaagagtcaccatt 245
Db 1306 RR 1247
Qy 246 gcatgggacacaaacacatcaacagagagccagccaccttggcttcaccgtcaa 305
Db 1246 RR 1187
Qy 306 ttggaagtgttcagagtcaccactgtcttcacggccagtcgttcacagaggaatgg 365
Db 1186 RR 1127
Qy 366 gaagaggtctcgaagaccatgtgctgtcggtgcaagtgttaataatggtgatga 425
Db 1126 RR 1067
Qy 426 ctggaagctaccagggtcgcatcaa 452
Db 1066 ATCGAAGCTCCCTGACCTGCAGCCA 1040

RESULT 4
US-08-596-291-1
Sequence 1, Application US/08596291
Patent No. 5821075
GENERAL INFORMATION:
APPLICANT: GONEZ, LEONEL JORGE
APPLICANT: SARAS, JAN
APPLICANT: CLAESSON-WELSH, LENA
APPLICANT: HELDIN, CARL-HENRIK
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA

ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/596,291
FILING DATE: 09-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: L0461/7000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
TELEX: 92-1742 EZEKIEL
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HOMO SAPIENS
FEATURE:
NAME/KEY: CDS
LOCATION: 78..7475
US-08-596-291-1

Query Match 6.5%; Score 31.4; DB 1; Length 8040;
Best Local Similarity 57.7%; Pred. No. 4;
Matches 56; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 170 acagcagagtggaattcacagcagcaccatcacagcgcgttaacagccacatcaaatgaga 229
Db 4978 ACACGCTGGGAGTGGAGAGATGACTTAGTCACAGCTCCACCAACATATCAAAATCCA 5037
Qy 230 tcaagagtcaccattgcatgggacacaaacacacat 266
Db 5038 CTTGGAGTTCAGCTTGTGATCAGACTCTAAGCAACAT 5074

RESULT 5
US-09-100-804-1
Sequence 1, Application US/09100804
Patent No. 6066472
GENERAL INFORMATION:
APPLICANT: GONEZ, LEONEL JORGE
APPLICANT: SARAS, JAN
APPLICANT: CLAESSON-WELSH, LENA
APPLICANT: HELDIN, CARL-HENRIK
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM: Floppy disk
MEDIUM TYPE: Floppy disk

CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/664,855
FILING DATE: 17-JUN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/170,294
FILING DATE: 30-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9114259.6

```

; FILING DATE: 02-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01187
; FILING DATE: 01-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 224452/SEE36438USTD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: GSTII-27 - FIGURE 2
;
; US-08-664-855-1

```

```

Query Match      6.3%; Score 30.8; DB 2; Length 954;
Best Local Similarity 52.3%; Pred. No. 2;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY  34 atccctgcgagaagtgctcgtgactggtggaatggaccacagtcgtgggtcccaaatg 153
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  500 ACCACCGCTGGTGGCTCGCGCCAGGAAAGCGCGGACACGACGACGACGACGATG 441

QY  154 accatggggctgaaacagagaggtgaattcacaggcacctacatacaacagcgtaaca 213
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  440 GCGATCGCGGGGGCTCAGCTGGTGGGCTCCACCTCCAGCCACAGCTCCACCATGCC 381

QY  214 gccacatacaa 223
    | | | | |
Db  380 GTCTGCTCCA 371

```

```

RESULT 10
; US-09-049-289-1/c
; Sequence 1, Application US/09049289
; Patent No. 6066456
; GENERAL INFORMATION:
; APPLICANT: BRIDGES, IAN G.
; APPLICANT: BRIGHT, SIMON W.J.
; APPLICANT: GREENLAND, ANDREW J.
; APPLICANT: HOLT, DAVID C.
; APPLICANT: JEPSON, IAN
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: PLANT-DERIVED ENZYME AND DNA SEQUENCES
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,289
; FILING DATE:
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,294
; FILING DATE: 30-DEC-1993
; APPLICATION NUMBER: WO PCT/GB92/01187
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9114259.6
; FILING DATE: 02-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 204218/SEE36438/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: GSTII-27 - FIGURE 2
;
; US-09-049-289-1

```

```

Query Match      6.3%; Score 30.8; DB 3; Length 954;
Best Local Similarity 52.3%; Pred. No. 2;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY  94 atccctgcgagaagtgctcgtgactggtggaatggaccacagtcgtgggtcccaaatg 153
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  500 ACCACCGCTGGTGGCTCGCGCCAGGAAAGCGCGGACACGACGACGACGATG 441

QY  154 accatggggctgtaaacagagaggtgaattcacaggcacctacatacaacagcgtaaca 213
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  440 GCGATCGCGGGGGCTCAGCTGGTGGGCTCCACCTCCAGCCACAGCTCCACCATGCC 381

QY  214 gccacatacaa 223
    | | | | |
Db  380 GTCTGCTCCA 371

```

```

RESULT 11
; US-09-022-238-1/c
; Sequence 1, Application US/09022238
; Patent No. 6177244
; GENERAL INFORMATION:
; APPLICANT: SYTKOWSKI, ARTHUR J. and YANG, MEIHENG
; TITLE OF INVENTION: A NOVEL NFG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,238
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/644,326
; FILING DATE: 10-MAY-1996

```

```

:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Maravic-Magovcivic, Ivana
: REGISTRATION NUMBER: P-43,338
: REFERENCE/DOCKET NUMBER: NER-262CP
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)742-4214
:
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 400 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: cDNA
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 3..398
:
: US-09-022-238-1

```

Query Match 6.3%; Score 30.6; DB 4; Length 400;
Best Local Similarity 56.4%; Pred. NO. 1.4;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 292 ggcttcacggtcaattggaagttttcagagtcacacactgtcttcacgggcccagtgcttc 351
Db 284 GACCTCGCAGTCCAGCGGTGTTTCTGGAAC TGAGGCGCTGTCTAC AAGCGCATTTGCTCTG 225
Qy 352 atagacaggaatgggaagagggtccctgaagaccatggct 392
Db 224 CTGGGCGAGACTCGGGCGGGAGGAGGATGAAGGCCCTGGGGCT 184

RESULT 12
US-07-675-211-1/c
; Sequence 1, Application US/07675211
; Patent No. 5219737

GENERAL INFORMATION:
APPLICANT: KAJIYAMA, NAOKI
APPLICANT: NAKANO, EIICHI
TITLE OF INVENTION: MUTANT LUCIFERASE OF A FIREFLY, MUTANT
TITLE OF INVENTION: LUCIFERASE GENES, NOVEL RECOMBINANT DNAs CONTAINING THE
TITLE OF INVENTION: GENES AND A METHOD OF PRODUCING MUTANT LUCIFERASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: N.Y.
COUNTRY: U.S.A
ZIP: 10036

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/675,211
FILING DATE: 19910326
CLASSIFICATION: 435

```

ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Luciola cruciata
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1644
US-07-675-211-1

```

Query Match 6.3%; Score 30.6; DB 1; Length 1644;
Best Local Similarity 62.3%; Pred. No. 3.1;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy	23	tcattgtttcttctctcttgcaactccctttgaaactctcttgcaacgaagaagtg	82
Db	436	TTACTGTTTTCTGCTACTGTGTATAACTTTATCTAAGCCTTTTGTAGAACTAAATACAATTG	377
Qy	83	atggaccagagatccct	99
Db	376	TTGGTTTAGAGATACCT	360

RESULT 13
US-07-903-047-1/c
; Sequence 1, Application US/07903047

```

: Patent No. 5329285
:
: GENERAL INFORMATION:
:
: APPLICANT: Kajiyama, Naoki
: APPLICANT: Nakano, Eiichi
: TITLE OF INVENTION: Thermos
: TITLE OF INVENTION: Thermos
: TITLE OF INVENTION: DNA, A
: TITLE OF INVENTION: Lucifer
:
: NUMBER OF SEQUENCES: 11
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Pennie & Edmon
: STREET: 1155 Avenue of th
:
: CITY: New York

```

```

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,047
; FILING DATE: 19920623
;

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1644 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cDNA to mRNA
US-07-0903-047-1

Query Match 6.3%; Score 30.6; DB 1; Length 1644;
Best Local Similarity 62.3%; Pred. No. 3.1;
Matches 48: Conservative 0; Mismatches 29; Indels 0

 \dot{Q}_2

Result No.	Score	Query ^a		Length	DB	ID	Description
		Match					
1	486	100.0	486	21	AAZ49865	Potato proteinase	
2	394.2	81.1	576	21	AAZ49855	pGEMav avidin cDNA	
3	389.4	80.1	459	20	AAZ80197	Chicken avidin enc	
4	389.4	80.1	604	18	AAR73194	Wld-type avidin D	
5	386.4	79.5	5177	20	AZ020997	Bovine scavenger r	
6	279.8	57.6	484	18	AAT68502	Barley alpha amyla	
7	94.4	19.4	324	21	AAZ49854	Potato proteinase	
8	36.6	7.5	1700	23	AAZ85309	DNA encoding novel	
9	34.8	7.2	2781	23	ABL08242	Drosophila melanog	


```

Db 104 cccgggacccctgcagaaagtgctcgtgactggaatggaccacgacgtcgggctcc 163
Qy 148 aacatgaccatcgggcgtgtgaacagcagaggtgaattcaccagggcacctacacacagcc 207
Db 164 aacatgaccatcgggcgtgtgaacagcagaggtgaattcaccagggcacctacacacagcc 223
Qy 208 gtaacagccacatcaaatgagatcaaaagagtcaccattgcatgggacacaaaacaccatc 267
Db 224 gtaacagccacatcaaatgagatcaaaagagtcaccattgcatgggacacaaaacaccatc 283
Qy 268 aacaagagagaccagcccaccttggcttcacagtcacagtcacagtcacagtcacagtcac 327
Db 284 aacaagagagaccagcccaccttggcttcacagtcacagtcacagtcacagtcacagtcac 343
Qy 328 actgtcttcacggccagtgcttcacagtcacagtcacagtcacagtcacagtcacagtcac 387
Db 344 actgtcttcacggccagtgcttcacagtcacagtcacagtcacagtcacagtcacagtcac 403
Qy 388 tggctgctcggtccaagtgttaatgacattggtgacactggaagagaggtcctgaagaccatg 447
Db 404 tggctgctcggtccaagtgttaatgacattggtgacactggaagagaggtcctgaagaccatg 463
Qy 448 atcaacatcttcaactgcctgcgcacacagagaaggagtga 486
Db 464 atcaacatcttcaactgcctgcgcacacagagaaggagtga 502

RESULT 3
AAX80197
ID AAX80197 standard; cDNA; 459 BP.
AC AAX80197;
XX
XX
Pt 20-AUG-1999 (first entry)
DE
DE Chicken avidin encoding cDNA.
KW Avidin; streptavidin; batroxobin; fibrinogen converting enzyme;
KW hybrid; fusion protein; sealant; surgery; reduce bleeding; fibrin; ss.
XX
OS Gallus gallus.
XX
XX WO929838-A1.
XX
XX 17-JUN-1999.
XX
XX 09-DEC-1998; 98WO-US26086.
XX
XX 09-DEC-1997; 97US-0067978.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Cedarholm-Williams SA;
XX
XX WPI; 1999-385599/32.
XX
XX P-FSDB; AAY17867.
XX
XX A fibrinogen-converting enzyme fusion protein
XX
XX Disclosure; Page 26-27; 35pp; English.
XX

```

The present invention describes a fibrinogen-converting enzyme fusion protein (FCE). The fusion protein is a multidomain protein comprising: (a) a FCE; and (b) a first member of a binding pair, that is linked to the FCE chain; (i) directly by bonds utilizing the N-terminal amino groups, the C-terminal carboxy groups or side-chain functionalities; (ii) via a bifunctional linkage moiety linking the groups or functionalities; or (iii) by the first member binding to the second member of the binding pair, where the second member of the binding pair is covalently attached to the first polypeptide chain. The FCE can be used in a method for producing fibrin. Fibrin is useful as a sealant in surgery to, e.g. reduce bleeding by sealing blood vessels, and tissues

CC that have been dissected either in surgery or through wounding. The CC fusion protein allows for the removal of the fibrinogen converting enzyme from the fibrin sealant preparation via the binding of streptavidin to a biotin solid support. The present sequence encodes CC chicken (Gallus gallus) avidin as given in the present invention. XX
SQ Sequence 459 BP; 118 A; 131 C; 120 G; 90 T; 0 other;

Query Match 80.1%; Score 389.4; DB 20; Length 459;
Best Local Similarity 98.5%; Pred. No. 5.1e-113;
Matches 393; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Qy 88 ccagagatccctgcagaaaagtgcctgactgactggaaaatggaccaacgactcgggctcc 147
Db 61 cccggcctctcgcagaaaagtgcctgactgactggaaaatggaccaacgactcgggctcc 120
Qy 148 aacatgaccatcgggcgtgtgaacagcagaggtgaattcaccagggcacctacacagcc 207
Db 121 aacatgaccatcgggcgtgtgaacagcagaggtgaattcaccagggcacctacacagcc 180
Qy 208 gtaacagccacatcaaatgagatcaaaagagtcaccattgcatgggacacaaaacaccatc 267
Db 181 gtaacagccacatcaaatgagatcaaaagagtcaccattgcatgggacacaaaacaccatc 240
Qy 268 aacaagagagaccagcccaccttggcttcacagtcacagtcacagtcacagtcacagtcac 327
Db 241 aacaagagagaccagcccaccttggcttcacagtcacagtcacagtcacagtcacagtcac 300
Qy 328 actgtcttcacggccagtgcttcacagtcacagtcacagtcacagtcacagtcacagtcac 387
Db 301 actgtcttcacggccagtgcttcacagtcacagtcacagtcacagtcacagtcacagtcac 360
Qy 388 tggctgctcggtccaagtgttaatgacattggtgacactggaagagaggtcctgaagaccatg 447
Db 361 tggctgctcggtccaagtgttaatgacattggtgacactggaagagaggtcctgaagaccatg 420
Qy 448 atcaacatcttcaactgcctgcgcacacagagaaggagtga 486
Db 421 atcaacatcttcaactgcctgcgcacacagagaaggagtga 459

```

```

RESULT 4
AAT73194
ID AAT73194 standard; DNA; 604 BP.
XX
XX AC
XX AAT73194;
XX
XX 27-APR-1998 (first entry)
XX
XX Wild-type avidin DNA.
XX
XX Streptavidin; biotin; anti-interference reagent; detection; muten;
XX avidin; non-specific binding; ss.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX CDS 44..502
XX /*tag= a
XX /*product= avidin
XX
XX sig_peptide 44..115
XX /*tag= b
XX mat_peptide 116..499
XX /*tag= c
XX /*product= avidin
XX
XX DE19637718-A1.
XX
XX 02-OCT-1997.
XX
XX 16-SEP-1996; 96DE-1037718.
XX

```


XX Example 2: Fig 1: 11lpp; English.
 XX
 CC The present sequence encodes potato proteinase inhibitor-I (PPI-I). This
 CC is used in the preparation of a binary vector designed to express a
 CC chimeric polypeptide comprising Avidin mature peptide, a plant-toxicous
 CC protein, fused to the PPI-I signal peptide. The binary vector is
 CC targeted to the vacuole by PPI-I signal sequence. Transformation of
 CC plant genome with the vector can produce pest resistance in plants,
 CC plant derived products and stored harvest material. Pests that can be
 CC controlled include, cotton bollworm, tropical army-worm, European corn
 CC -borer or red mite, tobacco horn worm, loopers, rice stem borer, porina,
 CC cutworms, diamondback moth, potato tuber moth, codling moth, Indian meal
 CC moth, gypsy moth, Argentine stem weevil, clover root weevil, grass
 CC -grubs, corn rootworm, rice and wheat weevils, mealworms, flour beetles,
 CC black field cricket, locusts, sawflies, Western flower thrips,
 CC Hessian flies or two-spotted mite.
 XX
 SQ Sequence 324 BP; 101 A; 53 C; 68 G; 102 T; 0 other;

Query Match 19.4%; Score 94.4; DB 21; Length 324;
 Best Local Similarity 99.0%; Pred. No. 9.9e-20;
 Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 atggagtaaaagttgtcaccatcattgtttttcttcttcttgcactccctttgaaact 60
 |||||
 DB 1 atggagtcaaagttgtcaccatcattgtttttcttcttcttgcactccctttgaaact 60
 |||||
 QY 61 ctcttgccagcaaaagaagtgtgacccagagata 96
 |||||
 DB 61 ctcttgccagcaaaagaagtgtgacccagagata 96
 |||||

RESULT 8
 AAS85309
 ID AAS85309 standard; cDNA; 1700 BP.
 XX
 AC
 XX
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #21113.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 PN
 XX
 XX 11-OCT-2001.
 PD
 XX
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX
 XX 31-MAR-2000; 2000US-0540217.
 PR
 XX 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB; ABC21112.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX
 XX Claim 1; SEQ ID No 21113; 103pp; English.
 XX
 QC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1700 BP; 469 A; 420 C; 427 G; 384 T; 0 other;

Query Match 7.5%; Score 36.6; DB 23; Length 1700;
 Best Local Similarity 52.3%; Pred. No. 0.45;
 Matches 81; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 200 acaccatcaacaaagaggacccacccttctgctccacgcgcacattggaggttttcag 319
 |||||
 DB 193 acactgtgaagaagaggtgcttctctctctctctctctctctctctctctctctctct 252
 |||||
 QY 320 agtcacacactgtcttcacgggcccagtccttcacagacaggaatgggagggaggtcctga 379
 |||||
 DB 253 aggcctcccccagcctgctgaactacagagacacagcgcattttggagcagtggtgg 312
 |||||
 QY 380 agacatgtgctgtctgctgaggttaagttaattgaac 414
 |||||
 DB 313 acatctgtggtctgcgcgtacacatgtgacagcc 347
 |||||

RESULT 9
 ABL08242
 ID ABL08242 standard; cDNA; 2781 BP.
 XX
 AC ABL08242;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 19208.
 DE
 XX
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 FA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI: 2001-656860/75.
 DR
 XX P-PSDB; ABB64139.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions - PT

XX

PS	Claim 1; SEQ ID NO 19208; 21pp + Sequence Listing; English.
PA	Claim 1; SEQ ID NO 19208; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 2781 BP; 760 A; 590 C; 696 G; 735 T; 0 other;

Query Match 7.28: Score 34.8: DB 23: Length 2781:

Best Local Similarity	50.6%;	Pred. No. 2.1;	Indels	0;
Matches	84;	Conservative	0;	Mismatches
			82;	Gaps
				0;

1. **Introduction**
 2. **Background**
 3. **Methodology**
 4. **Results**
 5. **Discussion**
 6. **Conclusion**
 7. **References**
 8. **Appendix**
 9. **Figure 1**
 10. **Figure 2**
 11. **Figure 3**
 12. **Figure 4**
 13. **Figure 5**
 14. **Figure 6**
 15. **Figure 7**
 16. **Figure 8**
 17. **Figure 9**
 18. **Figure 10**
 19. **Figure 11**
 20. **Figure 12**
 21. **Figure 13**
 22. **Figure 14**
 23. **Figure 15**
 24. **Figure 16**
 25. **Figure 17**
 26. **Figure 18**
 27. **Figure 19**
 28. **Figure 20**
 29. **Figure 21**
 30. **Figure 22**
 31. **Figure 23**
 32. **Figure 24**
 33. **Figure 25**
 34. **Figure 26**
 35. **Figure 27**
 36. **Figure 28**
 37. **Figure 29**
 38. **Figure 30**
 39. **Figure 31**
 40. **Figure 32**
 41. **Figure 33**
 42. **Figure 34**
 43. **Figure 35**
 44. **Figure 36**
 45. **Figure 37**
 46. **Figure 38**
 47. **Figure 39**
 48. **Figure 40**
 49. **Figure 41**
 50. **Figure 42**
 51. **Figure 43**
 52. **Figure 44**
 53. **Figure 45**
 54. **Figure 46**
 55. **Figure 47**
 56. **Figure 48**
 57. **Figure 49**
 58. **Figure 50**
 59. **Figure 51**
 60. **Figure 52**
 61. **Figure 53**
 62. **Figure 54**
 63. **Figure 55**
 64. **Figure 56**
 65. **Figure 57**
 66. **Figure 58**
 67. **Figure 59**
 68. **Figure 60**
 69. **Figure 61**
 70. **Figure 62**
 71. **Figure 63**
 72. **Figure 64**
 73. **Figure 65**
 74. **Figure 66**
 75. **Figure 67**
 76. **Figure 68**
 77. **Figure 69**
 78. **Figure 70**
 79. **Figure 71**
 80. **Figure 72**
 81. **Figure 73**
 82. **Figure 74**
 83. **Figure 75**
 84. **Figure 76**
 85. **Figure 77**
 86. **Figure 78**
 87. **Figure 79**
 88. **Figure 80**
 89. **Figure 81**
 90. **Figure 82**
 91. **Figure 83**
 92. **Figure 84**
 93. **Figure 85**
 94. **Figure 86**
 95. **Figure 87**
 96. **Figure 88**
 97. **Figure 89**
 98. **Figure 90**
 99. **Figure 91**
 100. **Figure 92**
 101. **Figure 93**
 102. **Figure 94**
 103. **Figure 95**
 104. **Figure 96**
 105. **Figure 97**
 106. **Figure 98**
 107. **Figure 99**
 108. **Figure 100**
 109. **Figure 101**
 110. **Figure 102**
 111. **Figure 103**
 112. **Figure 104**
 113. **Figure 105**
 114. **Figure 106**
 115. **Figure 107**
 116. **Figure 108**
 117. **Figure 109**
 118. **Figure 110**
 119. **Figure 111**
 120. **Figure 112**
 121. **Figure 113**
 122. **Figure 114**
 123. **Figure 115**
 124. **Figure 116**
 125. **Figure 117**
 126. **Figure 118**
 127. **Figure 119**
 128. **Figure 120**
 129. **Figure 121**
 130. **Figure 122**
 131. **Figure 123**
 132. **Figure 124**
 133. **Figure 125**
 134. **Figure 126**
 135. **Figure 127**
 136. **Figure 128**
 137. **Figure 129**
 138. **Figure 130**
 139. **Figure 131**
 140. **Figure 132**
 141. **Figure 133**
 142. **Figure 134**
 143. **Figure 135**
 144. **Figure 136**
 145. **Figure 137**
 146. **Figure 138**
 147. **Figure 139**
 148. **Figure 140**
 149. **Figure 141**
 150. **Figure 142**
 151. **Figure 143**
 152. **Figure 144**
 153. **Figure 145**
 154. **Figure 146**
 155. **Figure 147**
 156. **Figure 148**
 157. **Figure 149**
 158. **Figure 150**
 159. **Figure 151**
 160. **Figure 152**
 161. **Figure 153**
 162. **Figure 154**
 163. **Figure 155**
 164. **Figure 156**
 165. **Figure 157**
 166. **Figure 158**
 167. **Figure 159**
 168. **Figure 160**
 169. **Figure 161**
 170. **Figure 162**
 171. **Figure 163**
 172. **Figure 164**
 173. **Figure 165**
 174. **Figure 166**
 175. **Figure 167**
 176. **Figure 168**
 177. **Figure 169**
 178. **Figure 170**
 179. **Figure 171**
 180. **Figure 172**
 181. **Figure 173**
 182. **Figure 174**
 183. **Figure 175**
 184. **Figure 176**
 185. **Figure 177**
 186. **Figure 178**
 187. **Figure 179**
 188. **Figure 180**
 189. **Figure 181**
 190. **Figure 182**
 191. **Figure 183**
 192. **Figure 184**
 193. **Figure 185**
 194. **Figure 186**
 195. **Figure 187**
 196. **Figure 188**
 197. **Figure 189**
 198. **Figure 190**
 199. **Figure 191**
 200. **Figure 192**
 201. **Figure 193**
 202. **Figure 194**
 203. **Figure 195**
 204. **Figure 196**
 205. **Figure 197**
 206. **Figure 198**
 207. **Figure 199**
 208. **Figure 200**
 209. **Figure 201**
 210. **Figure 202**
 211. **Figure 203**
 212. **Figure 204**
 213. **Figure 205**
 214. **Figure 206**
 215. **Figure 207**
 216. **Figure 208**
 217. **Figure 209**

Qy 93 gatccctgccagaaagtgcctgcctgactgggaaatggaccaacgatctgggctccaacat 152

Db 124 gctgtcttcgaggcagtgaggcaatgatccgggcatccgacctactggcggttctggag 183

153

153 gaccatcggggcctgtgaacagcagaggtgaattcacaggcacctacatcacagccgtaac 212

D_b 184 aggaatcgcaagtgtgaacagaaacagtccaaaaacttctatcacaatccgcaatgacaaq 243

RESULT 10

ABL08255

ID ABL08255 standard; cDNA: 4512 BP.

AC ABL08255:

26-MAR-2002 (first entry)

DE *Protophila melanogaster* expressed polynucleotide SEQ ID NO 19247.

KW Drosophila: developmental biology: cell signalling: insecticide:

KW pharmaceutical; gene: ss

OS *Drosophila melanogaster*.

PN WO200171042-A2.

27-SEP-2001

23-MAR-2001: 2001WO-IIS09231

23-MAR-2000 2000US-191637P

PR 11-JUL-2000: 2000US-0614150.

PA (PEKE) PE CORP NY

PI Venter JC, Adams M, Li PWD, Myers FW:

AA
DB
WPT: 2001-656860/75

DR WFL, 2001-030000/
P-PSDB: ABB64152.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

PS
XX
Claim 1: SEO ID NO 19247: 21pp + Sequence Listing: English:

SQ sequence 2543 BP; 740 A; 477 C; 557 G; 769 T; 0 other;

	Query Match	6.9%	Score 33.4	DB 21	Length 2543
	Best Local Similarity	57.0%	Prod. No. 5.7		
	Matches	61	Conservative	0	Mismatches 46
					Indels 0
					Gaps 0
QY	202	acagccgttaacagccacatcaaatgaagatcaaaaagagtcacccattgcatgggagacacaaaac	261		
Db	2334	aaacagcaacgaacatcacagagacaaagggctcactgaccgcgacacacacaaac	2393		
QY	262	accatcaacaagaggaccagccacctttgtggttcaccggtcaattg	308		
Db	2394	aacacacaaacgtctcatctcaattctattcttctcaacaacattg	2440		

Search completed: September 20, 2002, 16:25:08
Job time: 3747 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 14:48:21 ; Search time 1842.77 Seconds
(without alignments)
5519.027 Million cell updates/sec

Title: US-09-743-690-6
Perfect score: 486
Sequence: 1 atggagtcacaaagtttgctca.....tgcgcacacagaaggagtga 486

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgtgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%	Score	Match	Length	DB	ID	Description
--------	-------	---	-------	-------	--------	----	----	-------------

1	389.4	80.1	604	5	GGAVIR	X05343	Chicken mRN
2	389.4	80.1	604	6	A93651	A93651	Sequence 3
3	389.4	80.1	604	6	AR177253	AR177253	Sequence
4	386.4	79.5	5177	6	AX018983	AX018983	Sequence
5	211.4	43.5	1119	11	CHKAV1D	L27818	Gallus gall
6	202.8	41.7	1224	5	GA311647	AJ311647	Gallus ga
7	164.8	33.9	277	5	GGAVR3MRN	Z21536	G.gallus av
8	156	32.1	277	5	GGAVR2MRN	Z21535	G.gallus av
9	130.2	26.8	1133	5	GGAVR3GEN	Z21612	G.gallus av
10	130.2	26.8	1133	5	GGAVR4X14	Z22883	G.gallus av
11	130.2	26.8	1133	5	GGAVR5X14	Z22882	G.gallus av
12	125.4	25.8	170	5	GGAVR1	Z97063	Gallus gall
13	118.4	24.4	1133	5	GA237659	AJ237659	Gallus ga
14	118.4	24.4	1133	5	GGAVR1GEN	Z21611	G.gallus av
15	116.8	24.0	1133	5	GA237658	AJ237658	Gallus ga
16	114.2	23.5	1335	5	GGAVR2GEN	Z21554	G.gallus av
17	114.2	23.5	1376	5	GA311648	AJ311648	Gallus ga
18	91.2	18.8	324	8	POTINH1X	L06137	Solanum tub
19	91.2	18.8	324	8	POTPIN1B	L06606	Solanum tub
20	91.2	18.8	324	8	STPIN1	X67675	S.tuberosum
21	91.2	18.8	324	8	STPINLMR	X67950	S.tuberosum
22	90.2	18.6	323	8	POTPINIA	L05985	Solanum tub
23	86.4	17.8	511	8	TOMWPII	K03290	Tomato leaf
24	78.4	16.1	480	8	TOMPII	J05094	L.peruvianu
25	75.2	15.5	435	8	STU30861	U30861	Solanum tub
26	48	9.9	2170	8	POTINH1	M17108	Potato woun
27	48	9.9	2506	8	TOMWPII	M19338	Tomato (L.e
28	46.4	9.5	1869	8	STPROINIX	Z14027	S.tuberosum
29	46.4	9.5	2437	8	STPROINI	Z12611	S.tuberosum
30	40.4	8.3	182760	2	AC094052	AC094052	Rattus no
31	40.2	8.3	2650	8	TOMTRYINH1	M59427	L.peruvianu
32	39.6	8.1	844	8	STPII	Z12819	S.tuberosum
33	39.4	8.1	447	8	NTPII3PI	Z12623	N.tabacum p
34	39.4	8.1	493	8	NTTIMPA	X67076	N.tabacum T
35	39.4	8.1	157774	2	AC013353	AC013353	Trypanoso
36	39.4	8.1	241098	2	AL662911	AL662911	Mus muscu
37	39	8.0	145385	2	AC091912	AC091912	Homo sapi
38	39	8.0	171280	2	AC024186	AC024186	Homo sapi
39	39	8.0	194409	2	AC106776	AC106776	Homo sapi
40	38.4	7.9	110258	2	AC104042	AC104042	Homo sapi
41	38.4	7.9	173119	2	AC079073	AC079073	Homo sapi
42	37.8	7.8	459	8	NTPII4PI	Z12619	N.tabacum P
43	37.8	7.8	479	8	NTTIMPB	X67075	N.tabacum T
44	37.4	7.7	125020	9	AF429315	AF429315	Homo sapi
45	37.2	7.7	3628	10	RNCADHER	X78997	R.norvegicu

ALIGNMENTS

RESULT	1	GGAVIR	604 bp	mRNA	linear	VRT 21-MAR-1995
LOCUS		GGAVIR				
DEFINITION		Chicken mRNA for avidin.				
ACCESSION		X05343				
VERSION		X05343.1	GI:63071			
KEYWORDS		avidin.				
SOURCE		chicken.				
ORGANISM		Gallus gallus				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS		Archosauria; Aves; Neognathae; Galliformes; Phasianidae;				
		Phasianinae; Gallus.				
		1 (bases 1 to 604)				
		Gope,M.L., Keinanen,R.A., Kristo,P.A., Conneely,O.M., Beattie,W.G.,				
		Zarucki-Schulz,T., O'Malley,B.W. and Kulomaa,M.S.				
TITLE		Molecular cloning of the chicken avidin cDNA				
JOURNAL		Nucleic Acids Res. 15 (8), 3595-3606 (1987)				
MEDLINE		87203384				
COMMENT		Data kindly reviewed (18-SEP-1987) by Kulomaa M.S.				
FEATURES		Location/Qualifiers				
		1..604				
		/organism="Gallus gallus"				
		/db_xref="taxon:9031"				

/tissue_type="oviduct (laying hen)"
/clone_lib="lambda gtl1"
29..34
RBS
/note="pot.rRNA binding site"
38..43
RBS
/note="pot.alternative rRNA binding site"
44..502
CDS
/note="precursor polypeptide (AA -24 to 128)"
/codon_start=1
/protein_id="CAA28954.1"
/db_xref="GI:63072"
/db_xref="SWISS-PROT:P02701"
/translation="MVHATSPHLLLLLSIALVAPGLSARKCSLTGKWTNDLGSNMTI
GAVNSGEFTQVITAVATPNSKEISPLHCTQNTINKRPTQPFQFVNWKFSESTTV
FTQCQFDIRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFRLRTOKE"
44..115
sig_peptide
116..499
mat_peptide
/product="mature avidin (AA 1-128)"
565..570
misc_feature
/note="pot.polyA signal"
594
polyA_site
/note="polyA site"
164 a 172 c 153 g 115 t
BASE COUNT
ORIGIN
Query Match 80.1%; Score 389.4; DB 5; Length 604;
Best Local Similarity 98.5%; Pred. No. 1e-109;
Matches 393; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 88 ccagagatccctgcagaaagtgcgtgactgggaatgguaacagatctgggtccc 147
Db 104 CCCGGCCTCTGCGCAAAAGTGTGCTGACTGGGAAATGGACCAAGATCTGGGCTCC 163
QY 148 aacatgaccatcggggtgctgactgggaatgguaacagatctgggtccc 207
Db 164 AACATGACCATCTGCGCAAAAGTGTGCTGACTGGGAAATGGACCAAGATCTGGGCTCC 223
QY 208 gtaacagcaccatcaaatgagatcaaatgagatgcacattgcatggagacacaaacacac 267
Db 224 GTACACGCCATCAATGAGATCAAGAGTCACCATGGATGGACACAAACACCATC 283
QY 268 aacaaggagagccagccaccccttggcttaccctgcaattgggaagtgttcagagtcacc 327
Db 284 AACAAAGAGAGCCAGCCACCCACCTTTGGCTTCACCGTCAATGGAAAGTTTCAGAGTCCACC 343
QY 328 actgtcttcacggggcagtgcttaccatagacaggaatgggaaggaggtctctgaagaccatc 387
Db 224 GTACACGCCATCAATGAGATCAAGAGTCACCATGGATGGACACAAACACCATC 283
QY 388 tggctgtcggtcagtgcttaccatagacaggaatgggaaggaggtctctgaagaccatc 447
Db 344 ACTGCTCTCAGCGGCCAGTGGTTCATAGACAGGAATGGGAAGGAGGTCTCTGAAGACCATG 403
QY 388 tggctgtcggtcagtgcttaccatagacaggaatgggaaggaggtctctgaagaccatc 447
Db 404 TGGCTGTCTCAGCGGCCAGTGGTTCATAGACAGGAATGGGAAGGAGGTCTCTGAAGACCATG 463
QY 448 atcaacatcttcactgcctgcgcacacagagaggtga 486
Db 464 ATCAACATCTTCACTCGCTGCGGCACACAGAGGAGTGA 502
RESULT 2
LOCUS A93651 604 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 3 from Patent EP0799890.
ACCESSION A93651
VERSION A93651.1 GI:6741840
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 604)
AUTHORS Mueller,R.D. and Deger,A.D.
TITLE Recombinant inactive core streptavidin mutants
JOURNAL Patent: EP 0799890-A 3 08-OCT-1997;

BOEHRINGER MANNHEIM GMBH (DE)
Location/Qualifiers
1..604
/organism="unidentified"
/db_xref="taxon:32644"
44..115
sig_peptide
44..502
CDS
/note="unnamed protein product"
/CODON_start=1
/protein_id="CAB69588.1"
/db_xref="GI:6741841"
/translation="MVHATSPHLLLLLSIALVAPGLSARKCSLTGKWTNDLGSNMTI
GAVNSGEFTQVITAVATPNSKEISPLHCTQNTINKRPTQPFQFVNWKFSESTTV
FTQCQFDIRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFRLRTOKE"
116..499
mat_peptide
/product="unnamed"
164 a 172 c 153 g 115 t
BASE COUNT
ORIGIN
Query Match 80.1%; Score 389.4; DB 6; Length 604;
Best Local Similarity 98.5%; Pred. No. 1e-109;
Matches 393; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 88 ccagagatccctgcagaaagtgcgtgactgggaatgguaacagatctgggtccc 147
Db 104 CCCGGCCTCTGCGCAAAAGTGTGCTGACTGGGAAATGGACCAAGATCTGGGCTCC 163
QY 148 aacatgaccatcggggtgctgactgggaatgguaacagatctgggtccc 207
Db 104 AACATGACCATCTGCGCAAAAGTGTGCTGACTGGGAAATGGACCAAGATCTGGGCTCC 223
QY 208 gtaacagcaccatcaaatgagatcaaatgagatgcacattgcatggagacacaaacacac 267
Db 224 GTACACGCCATCAATGAGATCAAGAGTCACCATGGATGGACACAAACACCATC 283
QY 268 aacaaggagagccagccaccccttggcttaccctgcaattgggaagtgttcagagtcacc 327
Db 284 AACAAAGAGAGCCAGCCACCCACCTTTGGCTTCACCGTCAATGGAAAGTTTCAGAGTCCACC 343
QY 328 actgtcttcacggggcagtgcttaccatagacaggaatgggaaggaggtctctgaagaccatc 387
Db 344 ACTGCTCTCAGCGGCCAGTGGTTCATAGACAGGAATGGGAAGGAGGTCTCTGAAGACCATG 403
QY 388 tggctgtcggtcagtgcttaccatagacaggaatgggaaggaggtctctgaagaccatc 447
Db 404 TGGCTGTCTCAGCGGCCAGTGGTTCATAGACAGGAATGGGAAGGAGGTCTCTGAAGACCATG 463
QY 448 atcaacatcttcactgcctgcgcacacagagaggtga 486
Db 464 ATCAACATCTTCACTCGCTGCGGCACACAGAGGAGTGA 502
RESULT 3
LOCUS ARI77253 604 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 3 from patent US 6312916.
ACCESSION ARI77253
VERSION ARI77253.1 GI:17919608
KEYWORDS
SOURCE Unknown.
ORGANISM UNKNOWN.
REFERENCE 1 (bases 1 to 604)
AUTHORS Kopetzki,E., Muller,R., Engh,R., Schmitt,U., Deger,A. and Brandstetter,H.
TITLE Recombinant inactive core streptavidin mutants
JOURNAL Patent: US 6312916-A 3 06-NOV-2001;
FEATURES Location/Qualifiers
1..604
source /organism="unknown"
BASE COUNT 164 a 172 c 153 g 115 t
ORIGIN

Query Match 80.1%; Score 389.4; DB 6; Length 604;
Best Local Similarity 98.5%; Pred. No. 1e-109;
Matches 393; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 88 ccagagatccctccagaaaagtctgcgtgactgggaatggaccaacgactctgggctcc 147
|||
Db 104 CCGGGCTCTCTGCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGGTCC 163
|||
QY 148 aacatgacatcgggctgtgaacagcagaggtgaattccacagcgaacctacatcacagcc 207
|||
Db 164 AACATGACCATCGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCC 223
|||
QY 208 gtaacagcacatcaaatgagatcaagagtcaccattgcatggagacacaaaacaccatc 267
|||
Db 224 GTACAGCCACATCAATGAGTCAAGAGTCACCACTGCTGGGACACAAACACCATC 283
|||
QY 268 aacaagaggaccagccacccttggcttcaccgtcaattggaagttttcagagtcacc 327
|||
Db 284 AACAGAGGACCCAGCCACCTTTGGCTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACC 343
|||
QY 328 actgtcttcacggccagtgcttcacatagacaggaatgggaaggttcctcgaagaccatg 387
|||
Db 344 ACTGTCTTCACGGCCAGTCTTCATAGACAGGAATGGGAAGAGGTCTCTGAAGACCATG 403
|||
QY 388 tggctgctcgggtcaagtgttaatgacattggtgactggaagctaccaggggtcggc 447
|||
Db 404 TGCTGCTGGGGTCAAGTGTAAATGACATTGGTGATGACTGGAAAGCTACCAAGGTCCGC 463
|||
QY 448 atcaacatcttactgcctgcgcacacagaaaggagtga 486
|||
Db 464 ATCAACATCTTACTCGCTCGCCGCACACAGAGGAGTGA 502
|||

RESULT 4
AX018983
LOCUS AX018983 5177 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9942577.
ACCESSION AX018983
VERSION AX018983.1 GI:10043078
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 5177)
AUTHORS Yla-Herttuala,S., Airenne,K., Kulomaa,M., Marjomaki,V. and Lehtolainen,P.
TITLE Biotin-binding receptor molecules
JOURNAL Patent: WO 9942577-A 1.26-AUG-1999;
YLA HERTTUALA SEPPO (FI); AIRENNE KARI (FI); KULOMAA MARKKU (FI); MARJOMAKI VARPU (FI); LEHTOLAINEN PAULIINA (FI); EUROGENE LIMITED (GB)
FEATURES Location/Qualifiers
source 1..5177
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Designed DNA sequence to encode transmembrane protein."
1071..2273
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07722.1"
/db_xref="GI:10043079"
/translation="NAQWDFDPQEDTDCSTESVKFDRSVTALLPHPKNGPTLQE RMKSKYKTAITLILYLVFVVLPIGIVAAQLLKWKNTVGSVNADISFSPGKNG SDENRFRFAVMRSMESRILQYLSDNAELLDAKNFQNSITTDQRFNDVLFLQNS LLSTQEHENITGDISKSLVGLNTVLDFQFSTETLNGRVOENAFKQOEEMKLEERI YNASAEIKSLDEQVYLEQEIKGEMKLLNITNDLRKDWHSQTLKNTLLQGARKC SLTGKWTNDLGSMTIGAVNSRGETGTIVATATSNKESPLGTQNTINKRTP TFGFTVNNKFSSTVFTGQCFTIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINF TRLRTRQKE"

BASE COUNT 1359 a 1221 c 1389 g 1208 t
ORIGIN

Query Match 79.5%; Score 386.4; DB 6; Length 5177;
Best Local Similarity 99.7%; Pred. No. 1.4e-108;
Matches 387; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 99 tccagaaagtgcgtgactgggaatggaccaacgactctgggctccaacatgacct 158
|||
Db 1886 TGCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAACATGACCAT 1945
|||
QY 159 cgggctgtgaacagcagaggtgaattcacagcgcacctacatcacagcgcgttaacagccac 218
|||
Db 1946 CGGGCTGTGAACACGACAGGTGAATTCACAGGCACCTACATCACAGCCGTAAACAGCCAC 2005
|||
QY 219 atcaaatgagatcaaaagagtcaccattgcatgggacacaaaacacacatcaacaagaggac 278
|||
Db 2006 ATCAAAATGAGATCAAAAGAGTCACCACTGCATGGGACACAAAACACCATCAACAAGAGGAC 2065
|||
QY 279 ccagccacccttggcttcaccgtcaattggaagttttcagagtcaccacactgtcttcac 338
|||
Db 2066 CCAGCCCACTTTGGCTTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACCACTGTCTTCAC 2125
|||
QY 339 gggccagtgctctacagacaggaatgggaagaggtcctcgaagacacatgtgctgctcg 398
|||
Db 2126 GGGCAGTGCTTCATAGACAGGAATGGGAAGAGGTCTCTGAAGACCATGTGCTGCTGCG 2185
|||
QY 399 gtcaagtgttaatgacattggtgactggaaagctaccagggttcgggcatcaacatctt 458
|||
Db 2186 GTCAAGTGTTAATGACATTGGTGATGACTTGGAAAGCTACCAAGGTTCGGCATCAACATCTT 2245
|||
QY 459 cactgcctgcgcacacagaaaggagtga 486
|||
Db 2246 CACTGCCTGCGCACACAGAGGAGTGA 2273
|||

RESULT 5
CHKAVID
LOCUS Gallus gallus (clone 44) STS, avidin gene, complete cds, sequence tagged site.
DEFINITION L27818
ACCESSION L27818.1 GI:450255
KEYWORDS STS; avidin; primer; sequence tagged site.
SOURCE Gallus gallus (strain White Leghorn, sub_species domesticus) female adult oviduct DNA.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (sites)
AUTHORS Delange,R.J. and Huang,T.-S.
TITLE Egg white avidin III. Sequence of the 78-residue middle cyanogen bromide peptide. Complete amino acid sequence of the protein subunit
JOURNAL J. Biol. Chem. 246, 698-709 (1971)
MEDLINE 71107558
REFERENCE 2 (sites)
AUTHORS Gope,M.L., Keinaenen,R.A., Kristo,P.A., Conneely,O.M., Beattie,W.G., Zarucki-Schulz,T., O'Malley,B.W. and Kulomaa,M.S.
TITLE Molecular cloning of the chicken avidin cDNA
MEDLINE Nucleic Acids Res. 15, 3595-3606 (1987)
REFERENCE 3 (bases 1 to 1119)
AUTHORS Wallen,M.J., Laukkanen,M.O. and Kulomaa,M.S.
TITLE Sequence of the chicken egg-white avidin gene
JOURNAL Unpublished (1994)
FEATURES Location/Qualifiers
source 1..1119
/organism="Gallus gallus"
/strain="White Leghorn"
/sub_species="domesticus"


```

Db 252 CTTGCTACAGTCTCGCTGACTGGGAATGGGACAAACGACCTGGGCTCAACATGACCA 311
QY 158 tcggggctgaaacagagagtggaattcacaggaacctacatcacagccgttaacagcca 217
Db 312 TCGGGCTGTGAACAGCAAGGTGAATTCACAGGCACCTACACCACGCGTAAACAGCCA 371
QY 218 catcaaatgagatcaagaagtcaccattgcatgggacacaaacacacatcaacaagagga 277
Db 372 CATCAATGAGATCAAGAGTCACCACTGCATGGGACACAAACACCATCAACAAGAGGA 431
QY 278 ccagcccaaccttggcttcacgcgtcaattggaagtttcaag 319
Db 432 CCCAGCCCACCTTTGGCTTCACTGTCAATTTGGAAGTTTTCAG 473

RESULT 7
GGAVR3MRN
LOCUS G.gallus avr3 mRNA for avidin related gene 3, partial. linear VRT 16-DEC-1997
DEFINITION G.gallus avr3 mRNA for avidin related gene 3, partial.
ACCESSION Z21536
VERSION Z21536.1 GI:65429
KEYWORDS avdin; avr3 gene.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 277)
AUTHORS Wallen, M.J.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-1993) Wallen M. J., University of Tampere,
Department of Biomedical Sciences, Tampere, Finland, SF-33101
REFERENCE 2 (bases 1 to 277)
AUTHORS Keinanen, R.A., Wallen, M.J., Kristo, P.A., Laukkanen, M.O.,
Toimela, T.A., Helenius, M.A. and Kulomaa, M.S.
TITLE Molecular cloning and nucleotide sequence of chicken avidin related
genes 1-5
JOURNAL Eur. J. Biochem. In press
REFERENCE 3 (bases 1 to 277)
AUTHORS Kunnas, T.A., Wallen, M.J. and Kulomaa, M.S.
TITLE Induction of chicken avidin and related mRNAs after bacterial
infection
JOURNAL Biochim. Biophys. Acta 1216 (3), 441-445 (1993)
MEDLINE 94092737
FEATURES
source
Location/Qualifiers
1..277
/organism="Gallus gallus"
/strain="White Leghorn"
/db_xref="taxon:9031"
/sex="Female"
/dev_stage="immature, estradiol treated"
/tissue_type="intestine"
/clone="pCAVR3-fl"
1..277
/partial
/gene="avr3"
/standard_name="Avidin related gene 3 mRNA"
/citation=[2]
1..277
/gene="avr3"
1
/gene="avr3"
/citation=[3]
/replace="t"

BASE COUNT 74 a 73 c 76 g 54 t
ORIGIN

Query Match 33.9%; Score 164.8; DB 5; Length 277;
Best Local Similarity 87.0%; Pred. No. 4.2e-40;
Matches 181; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 279 ccagcccaaccttggcttcacgcgtcaattggaagtttccagagtcaccactgtcttcac 338

```

```

Db 38 CCACCCACCTTTGGCTTCACTGTCCATTGGAACCTTTTCAGAGTCCACAGTGTCTTTGT 97
QY 339 gggccagtgcttcatacacaggaatggaagaggtcctgaagaccatgttgctctcgcg 398
Db 98 GGGCCAGTGCTTCATAGACAGGCGGAAAGGAGGTCTCTGAAGACCAATGGCTGCAACG 157
QY 399 gtcaagtgttaatacattgctgactgactggaagactaccagggtcgccatcaacatctt 458
Db 158 GTTAGCAGTTGATGACATTAGTACTGTAAGAGCTACCAAGGTCGGCTACACAACACTT 217
QY 459 cactcgctgcgcacacagaagagtgga 486
Db 218 CACTCGCCAGGCGCACAGTGGAGGAGTGA 245

RESULT 8
GGAVR2MRN
LOCUS G.gallus avr2 mRNA for avidin related gene 2, partial. linear VRT 03-FEB-1994
DEFINITION G.gallus avr2 mRNA for avidin related gene 2, partial.
ACCESSION Z21535
VERSION Z21535.1 GI:65428
KEYWORDS avdin; avr2 gene.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 277)
AUTHORS Kunnas, T.A., Wallen, M.J. and Kulomaa, M.S.
TITLE Induction of chicken avidin and related mRNAs after bacterial
infection
JOURNAL Biochim. Biophys. Acta 1216 (3), 441-445 (1993)
MEDLINE 94092737
FEATURES
source
Location/Qualifiers
1..277
/organism="Gallus gallus"
/strain="White Leghorn"
/db_xref="taxon:9031"
/sex="Female"
/dev_stage="immature, estradiol treated"
/tissue_type="oviduct"
/clone="pCAVR2-fl"
1..277
/partial
/gene="avr2"
/standard_name="Avidin related gene 2 mRNA"
/citation=[1]
1..277
/gene="avr2"

BASE COUNT 75 a 72 c 75 g 55 t
ORIGIN

Query Match 32.1%; Score 156; DB 5; Length 277;
Best Local Similarity 81.8%; Pred. No. 2.3e-37;
Matches 180; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 267 caacaagagagccagccacaccttggcttcacgcgtcaattggaagtttccagagtcacc 326
Db 26 CAAAGAGAGCCAGCCACCCACTTTGGCTTCACTGTCCATTTGGAACTTTTCAGAGTCCAC 85

```



```

MEDLINE      88260103
REMARK       (sites)
REFERENCE    2 (bases 1 to 1334)
AUTHORS      Keinanen,R.A., Wallen,M.J., Kristo,P.A., Laukkanen,M.O.,
              Toimela,T.A., Helenius,M.A. and Kulomaa,M.S.
TITLE        Molecular cloning and nucleotide sequence of chicken avidin-related
              genes 1-5
JOURNAL      Eur. J. Biochem. 220 (2), 615-621 (1994)
MEDLINE      94170814
REFERENCE    3 (bases 1 to 1334)
AUTHORS      Wallen,M.J.
TITLE        Direct Submission
JOURNAL      Submitted (04-JUN-1993) Wallen M. J., University of Tampere,
              Department of Biomedical Sciences, Tampere, Finland, SF-33101
FEATURES     Location/Qualifiers
              source
                1. .1334
                /organism="Gallus gallus"
                /strain="White Leghorn"
                /db_xref="taxon:9031"
                /clone="pgAV2H48"
                /sex="Female"
                /tissue_type="oviduct"
                /clone_lib="genomic library (EMBL4) of Bert W. O'Malley"
                /dev_stage="adult"
                join(<177..294,379..583,1009..1129,1217..>1334)
                /gene="avr4"
                /standard_name="Avidin related gene 4"
                /note="putative exons 1-4"
                /citation=[2]
                join(177..294,379..583,1009..1129,1217..1334)
                /gene="avr4"
BASE COUNT   259 a 440 c 344 g 291 t
ORIGIN
-
Query Match      26.8%; Score 130.2; DB 5; Length 1334;
Best Local Similarity 79.3%; Pred. No. 3.4e-29;
Matches 169; Conservative 0; Mismatches 38; Indels 6; Gaps 1;

QY 107 agtgcgtgactgggaatggaccacagatctgggtcccaacatgacctcggggctg 166
Db 377 AGTGCTCGCTGACTGGGAAATGGACCAACCTGGGCTCCATCAGCTGATCCAGCATCAGGCGTG 436

QY 167 tgaacagcagaggtgaattcacaggcacctacatcacgccgttaacagccacatcaaatg 226
Db 437 TGAACAGCAGAGGTGAATTTCACAGGACCTACCTCAGCGTGTAGCGGATTAATCCAGGAA 496

QY 227 agatcaaaagagtcaccattgcattgggacacaaacaccatcaacagagagagccagccca 286
Db 497 ACATCAGCTATCACCACCTGCTTGGGATCCACA-----CAAAAGAGCCAGCCGCCCA 550

QY 287 ccttggcttccaccgtcaattggaagtcttcag 319
Db 551 CCTTTGGCTTCACTGTCCATTTGGAACCTTTTCAG 583

RESULT 11
GGAVRSX14
LOCUS          G.gallus avr5 gene, exons 1-4.
DEFINITION    1334 bp
ACCESSION     222882
VERSION       222882.1 GI:311812
KEYWORDS      avridin; avr5 gene.
SOURCE        chicken.
ORGANISM      Gallus gallus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE     1 (bases 1 to 1334)
AUTHORS      Keinanen,R.A., Laukkanen,M.L. and Kulomaa,M.S.
TITLE        Molecular cloning of three structurally related genes for chicken
              avidin
JOURNAL      J. Steroid Biochem. 30 (1-6), 17-21 (1988)

```

```

MEDLINE      88260103
REMARK       (sites)
REFERENCE    2 (bases 1 to 1334)
AUTHORS      Keinanen,R.A., Wallen,M.J., Kristo,P.A., Laukkanen,M.O.,
              Toimela,T.A., Helenius,M.A. and Kulomaa,M.S.
TITLE        Molecular cloning and nucleotide sequence of chicken avidin-related
              genes 1-5
JOURNAL      Eur. J. Biochem. 220 (2), 615-621 (1994)
MEDLINE      94170814
REFERENCE    3 (bases 1 to 1334)
AUTHORS      Wallen,M.J.
TITLE        Direct Submission
JOURNAL      Submitted (04-JUN-1993) Wallen M. J., University of Tampere,
              Department of Biomedical Sciences, Tampere, Finland, SF-33101
FEATURES     Location/Qualifiers
              source
                1. .1334
                /organism="Gallus gallus"
                /strain="White Leghorn"
                /db_xref="taxon:9031"
                /clone="pgAV2H52"
                /sex="Female"
                /tissue_type="oviduct"
                /clone_lib="genomic library (EMBL4) of Bert W. O'Malley"
                /dev_stage="adult"
                join(<177..294,379..583,1009..1129,1217..>1334)
                /gene="avr5"
                /standard_name="Avidin related gene 5"
                /note="putative exons 1-4"
                /citation=[2]
                join(177..294,379..583,1009..1129,1217..1334)
                /gene="avr5"
BASE COUNT   260 a 440 c 343 g 291 t
ORIGIN
-
Query Match      26.8%; Score 130.2; DB 5; Length 1334;
Best Local Similarity 79.3%; Pred. No. 3.4e-29;
Matches 169; Conservative 0; Mismatches 38; Indels 6; Gaps 1;

QY 107 agtgcgtgactgggaatggaccacagatctgggtcccaacatgacctcggggctg 166
Db 377 AGTGCTCGCTGACTGGGAAATGGACCAACCTGGGCTCCATCAGCTGATCCAGCATCAGGCGTG 436

QY 167 tgaacagcagaggtgaattcacaggcacctacatcacgccgttaacagccacatcaaatg 226
Db 437 TGAACAGCAGAGGTGAATTTCACAGGACCTACCTCAGCGTGTAGCGGATTAATCCAGGAA 496

QY 227 agatcaaaagagtcaccattgcattgggacacaaacaccatcaacagagagagccagccca 286
Db 497 ACATCAGCTATCACCACCTGCTTGGGATCCACA-----CAAAAGAGCCAGCCGCCCA 550

QY 287 ccttggcttccaccgtcaattggaagtcttcag 319
Db 551 CCTTTGGCTTCACTGTCCATTTGGAACCTTTTCAG 583

RESULT 12
GGAVR1
LOCUS          Gallus gallus mRNA for avidin related protein 1, partial.
DEFINITION    170 bp
ACCESSION     297063
VERSION       297063.1 GI:2226407
KEYWORDS      avridin related protein; avr1 gene.
SOURCE        chicken.
ORGANISM      Gallus gallus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE     1 (bases 1 to 170)
AUTHORS      Keinanen,R.A., Wallen,M.J., Kristo,P.A., Laukkanen,M.O.,
              Toimela,T.A., Helenius,M.A. and Kulomaa,M.S.
TITLE        Molecular cloning and nucleotide sequence of chicken avidin-related
              genes 1-5

```


Db	167	CTTCGTGCACATGCTCGCTGACTGGGAATGGGGACAAACAACTGGGTCCCAACATGACCA	226
Qy	158	tcggggctgtgaacagcagagtgtaattccaggagcactcatcacacgcqtaacagcca	217
Db	227	TCGGAGCTGTGAACAGCAANTGGCAGTTCATGTCACCTACATCACAGCTGTACGACATA	286
Qy	218	catcaatgagatcaacagagtcaccattgcattgcatgggacacaaacaccatcaacaagagga	277
Db	287	ATCCAGGAAACATCAAGCTATCACCACTGCTGTGGATCCAAACA-----CAAAGAGCCT	340
Qy	278	ccagccacattgggttcacccgtcaattggaagtcttcaag	319
Db	341	GCCAGCCACCTTTGGCTTCACTGCTCCATTGGAACCTTTTCAG	382
RESULT 14			
GGAVRIGEN			
LOCUS	GGAVRIGEN	1335 bp	DNA linear VRT 18-MAR-2001
DEFINITION	G.gallus avr1 gene.		
ACCESSION	Z21611.1 Z21623		
VERSION	Z21611.1 GI:65431		
KEYWORDS	avidin; avr1 gene.		
SOURCE	chicken.		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae;		
	Phasianinae; Gallus.		
	1 (bases 1 to 1335)		
AUTHORS	Keinanen,R.A., Laukkanen,M.L. and Kulomaa,M.S.		
	Molecular cloning of three structurally related genes for chicken		
	avidin		
	J. Steroid Biochem. 30 (1-6), 17-21 (1988)		
JOURNAL	J. Steroid Biochem. 30 (1-6), 17-21 (1988)		
MEDLINE	88260103		
REMARK	(sites)		
REFERENCE	2 (bases 1 to 1335)		
	Keinanen,R.A., Wallen,M.J., Kristo,P.A., Laukkanen,M.O.,		
	Toimela,T.A., Helenius,M.A. and Kulomaa,M.S.		
	Molecular cloning and nucleotide sequence of chicken avidin-related		
AUTHORS	genes 1-5		
	Eur. J. Biochem. 220 (2), 615-621 (1994)		
	94170814		
	3 (bases 1 to 1335)		
TITLE	Wallen,M.J.		
	Direct Submission		
	Submitted (01-FEB-1993) Wallen M. J., University of Tampere,		
	Department of Biomedical Sciences, Tampere, Finland, SF-33101		
JOURNAL	Location/Qualifiers		
	1. .1335		
	/organism="Gallus gallus"		
	/strain="White Leghorn"		
FEATURES	/db_xref="taxon:9031"		
	/clone="pgAV1.8"		
	/sex="Female"		
	/tissue_type="oviduct"		
source	/clone_lib="genomic library (EMBL4) of Bert W. O'Malley"		
	/dev_stage="adult"		
	join(<177..294,379..583,1010..1130,1218..>1335)		
	/gene="avr1"		
mRNA	/standard_name="Avidin related gene 1"		
exon	/note="putative exons 1-4"		
	/citation=[1]		
	<177..294		
	/gene="avr1"		
gene	/number=1		
	join(177..294,379..583,1010..1130,1218..1335)		
	/gene="avr1"		
	379..583		
exon	/gene="avr1"		
	/number=2		
	1010..1130		
	/gene="avr1"		
exon	/number=3		
	1218..>1335		

```

/product="avidin-related protein 6 (AVR6)"
/protein_id="CAB39893.1"
/db_xref="GI:4539028"
/translation="MVHATSPDLLLLLSLALVAPGLSARKCSLTGEWDNNLGSIMTI
GAVNDGEFNGTAVTAVADPCNTKIGPLICIOHKRACDPFCPTVHNNFSESISVFF
GOCFVDRSGKEYLTKRWLQLRLAYDDISDDPHKATRYGYNNTFRQRTVEE"
<13..84

```

sig_peptide

```

/join(85..93,178..382,808..928,1016..1110)
/gene="avr6"

```

mat_peptide

```

/product="avidin-related protein 6 (AVR6)"

```

intron

```

/number=1

```

exon

```

/number=2

```

intron

```

/number=3

```

exon

```

/number=4

```

intron

```

/number=5

```

exon

```

/number=6

```

polyA_signal

```

/number=7

```

```

BASE COUNT 201 a 388 c 293 g 251 t
ORIGIN

```

```

Query Match 24.0%; Score 116.8; DB 5; Length 1133;
Best Local Similarity 73.9%; Pred. No. 4.7e-25;
Matches 164; Conservative 0; Mismatches 52; Indels 6; Gaps 1;

Qy 98 CTGCCAGAAAGTCTCTGCTGACTGGGAAATGGACCAAGATCTGGGCTCCATGACCA 157
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 CTTCTCACAGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 226
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 158 TCGGGGCTGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 217
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 227 TCGGGGCTGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 286
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 218 CATCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 277
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 287 ATCCAGGAACATCAAGCTATCAAGCTATCAAGCTATCAAGCTATCAAGCTATCAAG 340
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 278 CCCAGCCACCTTGGCTTCAAGCTATCAAGCTATCAAGCTATCAAGCTATCAAGCT 319
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 341 GCCAGCCACCTTGGCTTCAAGCTATCAAGCTATCAAGCTATCAAGCTATCAAGCT 382

```

Search completed: September 20, 2002, 16:20:58
Job time: 5557 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:41:14 ; Search time 27.21 Seconds
(without alignments)
1023.601 Million cell updates/sec

Title: US-09-743-690-7

Perfect score: 845

Sequence: 1 MESKFAHIIIVFFLLATPFET.....DWKATRVGINIFRLRTQKE 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	670	79.3	152	13 Q98SH4	Q98sh4 gallus gall
2	151	17.9	107	10 Q07459	Q07459 solanum tub
3	151	17.9	107	10 Q43648	Q43648 solanum tub
4	146	17.3	106	10 Q43651	Q43651 solanum tub
5	136	16.1	529	5 Q25058	Q25058 heliocidari
6	129	15.3	406	5 Q25059	Q25059 heliocidari
7	122	14.4	101	10 Q41434	Q41434 solanum tub
8	97	11.5	30	10 Q41490	Q41490 solanum tub
9	95	11.2	1395	8 Q35059	Q35059 marchantia
10	84.5	10.0	400	5 Q9VQY1	Q9vqy1 drosophila
11	84.5	10.0	749	5 Q967D9	Q967d9 drosophila
12	84.5	10.0	903	5 Q967D8	Q967d8 drosophila
13	84.5	10.0	1531	5 Q967D7	Q967d7 drosophila
14	82.5	9.8	1153	10 Q64728	Q64728 arabidopsis
15	82.5	9.8	1218	10 Q9LKQ4	Q9lkq4 arabidopsis
16	81	9.6	570	3 Q9HFF8	Q9hff8 saccharomyc

17	80	9.5	1592	10	Q91G74	Q91gt4 oryza sativ
18	78.5	9.3	445	5	Q9W0F6	Q9w0f6 drosophila
19	78.5	9.3	802	5	P91151	P91151 caenorhabdi
20	77.5	9.2	434	11	Q35317	Q35317 mus musculu
21	77.5	9.2	475	11	O62664	O62664 rattus norv
22	77.5	9.2	1124	4	O60280	O60280 homo sapien
23	77	9.1	401	6	Q08835	Q08835 cercopitheci
24	75.5	8.9	263	3	O13854	O13854 schizosacch
25	75.5	8.9	539	2	Q9RKQ8	Q9rkq8 streptomyc
26	75	8.9	465	10	Q9CAN8	Q9can8 arabidopsis
27	75	8.9	693	17	O58898	O58898 pyrococcus
28	75	8.9	1441	9	Q9AZL2	Q9azl2 bacterioph
29	75	8.9	1441	16	Q9CI85	Q9ci85 lactococcus
30	75	8.9	3201	5	Q9W0U2	Q9w0u2 drosophila
31	74.5	8.8	974	5	O01574	O01574 caenorhabdi
32	74	8.8	398	2	Q9RF12	Q9rf12 clostridium
33	74	8.8	424	16	Q9I456	Q9i456 pseudomonas
34	74	8.8	527	5	Q27471	Q27471 caenorhabdi
35	74	8.8	684	4	Q9UF31	Q9uf31 homo sapien
36	74	8.8	3542	5	Q9U5M2	Q9u5m2 plasmodium
37	73.5	8.7	253	4	Q96PG0	Q96pg0 homo sapien
38	73.5	8.7	531	10	Q9SZF7	Q9szf7 arabidopsis
39	73.5	8.7	725	16	Q9CFY6	Q9cfy6 lactococcus
40	73.5	8.7	1250	5	Q23409	Q23409 caenorhabdi
41	73	8.6	336	5	Q19962	Q19962 caenorhabdi
42	73	8.6	480	10	Q944R8	Q944r8 arabidopsis
43	73	8.6	632	10	Q9SM30	Q9sm30 taraxacum o
44	72.5	8.6	460	10	Q9LIF8	Q9lif8 arabidopsis
45	72.5	8.6	481	16	Q9RWA3	Q9rwa3 deinococcus

ALIGNMENTS

RESULT 1

Q98SH4 ID Q98SH4 PRELIMINARY; PRT; 152 AA.
AC Q98SH4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE AVIDIN PRECURSOR.
GN AVID.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ahlroth M.K.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21070478; PubMed=11167523;
RA Ahlroth M.K., Koia E.H., Ewald D., Masabanda J., Sazanov A., Fries R.,
RA Kulomaa M.S.;
RT "Characterization and Chromosomal Localization of the Chicken Avidin
Gene Family.";
RL Anim. Genet. 31:367-375(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95394357; PubMed=7665080;
RA Wallen M.J., Laakkonen M.O., Kulomaa M.S.;
RT "Cloning and sequencing of the chicken egg-white avidin-encoding gene
and its relationship with the avidin-related genes Avrl-Avr5.";
RL Gene 161:205-209(1995).
RL EMBL; AJ311647; CAC34569.1; -.
DR HSP; P02701; IAVD.
DR InterPro; IPR000086; Avidin.
DR Pfam; PF01382; Avidin; 1.
DR ProDom; PD016055; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.

```

KW Signal.
FT SIGNAL 1 24 POTENTIAL.
SQ SEQUENCE 152 AA: 16742 MW: 04B4965B070B974B CRC64:

Query Match 79.3%; Score 670; DB 13; Length 152;
Best Local Similarity 95.5%; Pred. No. 3.2e-57;
Matches 126; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 30 PEIPARKSLTGKWTNDLGSNMTICAVNSRGEFTGTYITAVTATSNKESPLHGTONVI 89
   I : |||||
DB 21 PGLSARKSLTGKWDNDLGSNMTIGAVNSKGEFTGTYITAVTATSNKESPLHGTONVI 80
   I : |||||
QY 90 NKRPQPTFGFVNNKFSSTVFTGQCPIDRNGKEVLKTMLLRSSVNDIGDDWKATRVG 149
   I : |||||
DB 81 NKRPQPTFGFVNNKFSSTVFTGQCPIDRNGKEVLKTMLLRSSVNDIGDDWKATRVG 140
   I : |||||
QY 150 INIPTRLTQKE 161
   I : |||||
DB 141 INIPTRLTQKE 152
   I : |||||

RESULT 2
ID Q07459 PRELIMINARY; PRT; 107 AA.
AC Q07459;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEINASE INHIBITOR I.
GN PINI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ILAM HARDY; TISSUE=POTATO TUBER;
RX MEDLINE=95106304; PubMed=7807552;
RA Reuning L.L., Spriggs T.W., Christeller J.T.;
RT "Evolution of the proteinase inhibitor I family and apparent lack of
RL Plant Physiol. 102:1061-1061(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ILAM HARDY;
RX MEDLINE=95106304; PubMed=7807552;
RA Reuning L.L., Spriggs T.W., Christeller J.T.;
RT "Evolution of the proteinase inhibitor I family and apparent lack of
RL J. Mol. Evol. 39:644-654(1994).
DR EMBL; X67950; CAA48136.1; -
DR EMBL; X67675; CAA47907.1; -
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit. 1.
DR PRINTS; PR00292; POTATOINHBT.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 107 AA; 12148 MW; DAB99CF6F11D9811 CRC64;

Query Match 17.9%; Score 151; DB 10; Length 107;
Best Local Similarity 93.8%; Pred. No. 3.2e-07;
Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESKFAHIIVFFLLATPFETLLARKESDGPEI 32
   I : |||||
DB 1 MESKFAHIIVFFLLATSFETLLARKESDGPEV 32
   I : |||||

RESULT 3
ID Q07459 PRELIMINARY; PRT; 107 AA.
AC Q07459;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEINASE INHIBITOR I.
GN PINI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ILAM HARDY; TISSUE=TUBER;
RX MEDLINE=95106304; PubMed=7807552;
RA Reuning L.L., Spriggs T.W., Christeller J.T.;
RT "Evolution of the proteinase inhibitor I family and apparent lack of
RL J. Mol. Evol. 39:644-654(1994).
DR EMBL; X67950; CAA48136.1; -
DR EMBL; X67675; CAA47907.1; -
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHBT.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 107 AA; 12148 MW; DAB99CF6F11D9811 CRC64;

Query Match 17.9%; Score 151; DB 10; Length 107;
Best Local Similarity 93.8%; Pred. No. 3.2e-07;
Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESKFAHIIVFFLLATPFETLLARKESDGPEI 32
   I : |||||
DB 1 MESKFAHIIVFFLLATSFETLLARKESDGPEV 32
   I : |||||

RESULT 3

```

```

Q43648
ID Q43648 PRELIMINARY; PRT; 107 AA.
AC Q43648;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEINASE INHIBITOR I.
GN PINI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ILAM HARDY; TISSUE=POTATO TUBER;
RX MEDLINE=95106304; PubMed=7807552;
RA Reuning L.L., Spriggs T.W., Christeller J.T.;
RT "Evolution of the proteinase inhibitor I family and apparent lack of
RL J. Mol. Evol. 39:644-654(1994).
DR EMBL; L06137; AAA69780.1; -
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHBT.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 107 AA; 12178 MW; DAADCDB3F11D9811 CRC64;

Query Match 17.9%; Score 151; DB 10; Length 107;
Best Local Similarity 93.8%; Pred. No. 3.2e-07;
Matches 30; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESKFAHIIVFFLLATPFETLLARKESDGPEI 32
   I : |||||
DB 1 MESKFAHIIVFFLLATSFETLLARKESDGPEV 32
   I : |||||

RESULT 4
ID Q43651 PRELIMINARY; PRT; 106 AA.
AC Q43651;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEINASE INHIBITOR I (FRAGMENT).
GN PINI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ILAM HARDY; TISSUE=TUBER;
RX MEDLINE=95106304; PubMed=7807552;
RA Reuning L.L., Spriggs T.W., Christeller J.T.;
RT "Evolution of the proteinase inhibitor I family and apparent lack of
RL J. Mol. Evol. 39:644-654(1994).
DR EMBL; L06985; AAA72133.1; -
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHBT.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
FT NON_TER 1
SQ SEQUENCE 106 AA; 12075 MW; OCB4A72602D2B4E2 CRC64;

```

DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
DE FIBROPELLIN III (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
OC Helicoidaris. OC

[1]
SEQUENCE FROM N.A.
RP
RA Bisgrove B.W.;
RT "Evolution of the fibropellin gene family and pattern of fibropellin
RT gene expression in sea urchin phylogeny";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L33862; AAA29996.1; -.
DR HSP; P00740; IEDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000088; Avidin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR000083; fibronectin_type_1.
DR InterPro; IPR002049; Laminin_EGF.

```

DR Pfam; PF01382; Avidin; 1.
DR Pfam; PF00008; EGF; 7.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFLAMININ.
DR PRINTS; PR00012; FNTYPEI.
DR PRODOM; PD016055; Avidin; 1.
DR SMART; SM00179; EGF_CA; 7.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_7.
DR PROSITE; PS00577; AVIDIN; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_7.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS01187; EGF_CA; 5.
DR KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
DR FT NON_TER 1
DR SQ SEQUENCE 406 AA; 43475 MW; EC89CCE8ABB7E89A CRC64;

Query Match 15.3%; Score 129; DB 5; Length 406;
Best Local Similarity 32.8%; Pred. No. 0.00022;
Matches 41; Conservative 20; Mismatches 48; Indels 16; Gaps

Qy 37 CSLTGKWTNGLGSNMNTIGAVNSRGFEFTGYITAVTATSNIEKESPL-HGTQNTI----NK 91
| | | | | : | | | | | : | | | | | : | | | | | : |
Db 281 CDLEGVWVNECDQITI-SKTS TGMMGLDHLT-----FNERLIGYGAPTVWVGYSK 331

Qy 92 RTQ-PTFGFTVNNKFSESTVPTGCFIDRNGKEVLKTMWLLRSVNDIGDDWKATRVGI 150
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 HVKPSFGFTVVRNGRSTTSMWTCGOCHLCA-GOEVLTYTWTIESNMVDTCEDIKKANKVGO 390

```

QY	151	NIFTR	155	
		:	:	!!
Db	391	DKWTR	395	
RESULT	7			
Q411434				
ID	Q411434	PRELIMINARY;	PRT;	101 AA.
AC	Q411434;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	WOUND-INDUCIBLE PROTEINASE INHIBITOR I (FRAGMENT).			
OS	Solanum tuberosum (potato).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asterales; Solanaceae; Solanales; Solanaceae; Solanum.			
NCBI_TaxID=4113;				
FN	!!			

RP SEQUENCE FROM N.A.
 RC STRAIN-KESZTHELYI 955 (WHITE LADY); TISSUE=ETIOLATED SHOOT;
 RX MEDLINE=96198594; PubMed=8925888;
 RA Banfalvi Z., Molnar A., Molnar G., Lakatos L., Szabo L.
 RT "Starch synthesis, and tuber storage protein genes are differently
 RL expressed in Solanum tuberosum and in Solanum brevidens.";
 DR EMBL: U30861; AAC49603.1; -;
 DR HSSP: F19873; IHXM.
 DR InterPro: IPR000864; Potato_inhibit.
 DR Pfam: PF00280; potato_inhibit.1.
 DR PRINTS: PR00292; POTATOINHBT.
 DR PRODOM: PD002604; Potato_inhibit.1.
 DR PROSITE: PS00285; POTATO_INHIBITOR; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 101 AA; 11396 MW; A27EG5CADE4D1CEA CRC64;

Query Match 14.4%; Score 122; DB 10; Length 101;
 Best Local Similarity 92.3%; Pred. NO. 0.00019;
 Matches 24; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 HIIIVFFLLATPEFTLLARKESDGPEI 32
 DB 1 HIIIVFFLLATPEFTLLARKESDGPEV 26
 |||||

RESULT 8
 Q41490 PRELIMINARY; PRT; 30 AA.
 ID Q41490
 AC Q41490; (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROTEINASE INHIBITOR I (FRAGMENT).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_taxid=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. RUSSET BURBANK;
 RA Lee J.S., Yang Y.S.;
 RT "Nucleotide sequence of 5' flanking region of potato proteinase
 RT inhibitor I gene.";
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z14027; CAA78402.1; -;
 FT NON_TER 30
 SQ SEQUENCE 30 AA; 3439 MW; 6090AB9BDA387D40 CRC64;

Query Match 11.5%; Score 97; DB 10; Length 30;
 Best Local Similarity 76.9%; Pred. NO. 0.011;
 Matches 20; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MESKFAHIIIVFFLLATPEFTLLARKE 26
 DB 1 MEAKFAHIIIVFFLLATPEFTFNQSE 26
 |||||

RESULT 9
 Q35059 PRELIMINARY; PRT; 1395 AA.
 ID Q35059
 AC Q35059; (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE TRNA-GLY PROTEIN.
 GN TRNA-GLY.
 OS Marchantia polymorpha (Liverwort).
 OC Mitochondrion.
 OX Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;

OC Marchantiales; Marchantiaceae; Marchantia.
 OX NCBI_taxid=3197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=921114051; PubMed=17311062;
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
 RA Kohchi T., Ogura Y., Kanegae T., Akashi K., Ohyama K.,
 RT "Gene organization deduced from the complete sequence of liverwort
 RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
 RT mitochondrial genome.";
 RL J. Mol. Biol. 223:1-7(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92319654; PubMed=1620617;
 RA Takemura M., Oda K., Yamato K., Ohta E., Nakamura Y., Nozato N.,
 RA Akashi K., Ohyama K.;
 RT "Gene clusters for ribosomal proteins in the mitochondrial genome of a
 RT liverwort, Marchantia polymorpha.";
 RL Nucleic Acids Res. 20:3199-3205(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92350681; PubMed=1641342;
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
 RA Akashi K., Ohyama K.;
 RT "Transfer RNA genes in the mitochondrial genome from a liverwort,
 RT Marchantia polymorpha: the absence of chloroplast-like tRNAs.";
 RL Nucleic Acids Res. 20:3773-3777(1992).
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
 CC GRADIENT ACROSS THE MEMBRANE. THE ALPHA CHAIN IS A REGULATORY
 CC SUBUNIT (BY SIMILARITY).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
 DR EMBL: M68929; AAC09444.1; -;
 DR InterPro: IPR000100; ATP-synt_ab_N.
 DR InterPro: IPR000194; ATPase_alpha_beta.
 DR InterPro: IPR000442; Intron_maturase2.
 DR InterPro: IPR000477; RVTse.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00006; ATP-synt_ab; 1.
 DR Pfam: PF02874; ATP-synt_ab_N; 1.
 DR Pfam: PF01348; Intron_maturase2; 1.
 DR Pfam: PF00078; rvt; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR ATP synthesis: ATP-binding; CF(1); Hydrogen ion transport; Hydrolase;
 KW Mitochondrion; RNA-directed DNA polymerase.
 SQ SEQUENCE 1395 AA; 156997 MW; 33DD91CF184F9686 CRC64;

Query Match 11.2%; Score 95; DB 8; Length 1395;
 Best Local Similarity 24.8%; Pred. NO. 1.9;
 Matches 41; Conservative 23; Mismatches 71; Indels 30; Gaps 7;

OY 10 VFFLLATPEFTLLARKESDG----PEIDARKCSLTGKWTNDLGSNMTIGAVNSRGFTGT 65
 DB 300 VFYLSRLLEAAKMSDQTGAGSUTALPVECAPTGR---SHIQWYGVGEVAGRYQT 356
 |||||

OY 66 YITAVTATSNKESPLHGTQNTINKRTQPTGPTVNM-----KFSEST-----TV 111
 DB 357 RSNPKDQSAARRIGRADIHGVOKPIKPSDPVNL-VOWNAASSKYDESSGKHPYKPI 415
 |||||

OY 112 FTGOCFIDRNGKEVLKTMWLR-SSVNDIGDDWKATRYGINFTR 155
 DB 416 HSGSAHLTRQGR-----LLRPSGVVTEGDKPKAYALCTRNYSK 453
 |||||

RESULT 10
 Q9VQY1 PRELIMINARY; PRT; 400 AA.
 ID Q9VQY1
 AC Q9VQY1;

DR InterPro; IPR003959; AAA_subfam.
DR InterPro; IPR00862; RFC.
DR Pfam; PF00004; AAA; 1.
SQ SEQUENCE 1218 AA; 135303 MW;

Query Match	9.8%	Score 82.5;	DB 10;	Length 1218;
Best Local Similarity	24.8%	Pred. No. 26;		
Matches	30;	Conservative 23;	Mismatches 47;	Indels 21;
Gaps	6;			

Search completed: September 17, 2002, 15:43:22
Job time: 128 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:41:34 ; Search time 11.91 Seconds
(without alignments)
523.413 Million cell updates/sec

Title: US-09-743-690-7

Perfect score: 845

Sequence: 1 MESKFAHIIIVFFLLATPFET.....DWKATRVGINIFRLRTQKE 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	681	80.6	152	1	AVID_CHICK
2	499	59.1	150	1	AVR4_CHICK
3	469	55.3	150	1	AVR3_CHICK
4	467	55.3	150	1	AVR7_CHICK
5	458	54.2	150	1	AVR1_CHICK
6	457	54.1	150	1	AVR6_CHICK
7	444	52.5	150	1	AVR2_CHICK
8	154	18.2	107	1	IC1D_SOLTU
9	143	16.9	107	1	IC1I_SOLTU
10	143	16.9	111	1	IC1I_LYCES
11	142.5	16.9	183	1	SAV1_STRVL
12	142.5	16.9	183	1	SAV_STRAV
13	140	16.6	1064	1	FBP1_STRPU
14	139.5	16.5	183	1	SAV2_STRVL
15	136	16.1	111	1	IC1I_LYCPE
16	124	14.7	570	1	FBP3_STRPU
17	85	10.1	481	1	UGDH_CAEEL
18	80.5	9.5	701	1	EPG_MICLU
19	77.5	9.2	434	1	PBX3_HUMAN
20	74.5	8.8	1268	1	PGCN_MOUSE
21	74	8.8	872	1	GUXA_CELFI
22	73.5	8.7	924	1	KE2_CANAL
23	72	8.5	331	1	TRPD_BUCSC
24	71.5	8.5	430	1	PBX1_HUMAN
25	71.5	8.5	729	1	PET2_HUMAN
26	70.5	8.3	630	1	MUC1_MOUSE
27	69.5	8.2	542	1	FIBP_ADECT
28	69.5	8.2	713	1	CDG2_PAEWA
29	69.5	8.2	1286	1	RPD_GUITH
30	68.5	8.1	292	1	MURE_BUCAP
31	68.5	8.1	435	1	VGLM_VZVD
32	68.5	8.1	535	1	ICAI_BOVIN
33	68	8.0	332	1	IPAD_SHIDY
					Q03947 shigella dy
					Q02496 mus musculus
					Q0426 homo sapien
					Q16348 homo sapien
					Q0426 homo sapien
					Q02496 mus musculus
					Q65914 canine aden
					P31835 paenibacill
					Q78483 guillardia
					O85298 buchnera ap
					P09298 varicella-z
					Q95132 bos taurus
					Q03947 shigella dy

RESULT 1

ID	AVID_CHICK	STANDARD;	PRT;	152 AA.
AC	P02701; Q91958;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Avidin precursor.			
GN	AVD.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87203384; PubMed=3575102;			
RA	Gope M.L., Keinaenen R.A., Kristo P.A., Conneely O.M., Beattie W.G.,			
RA	Zarucki-Schulz T., O'Malley B.W., Kulomaa M.S.;			
RT	"Molecular cloning of the chicken avidin cDNA.;"			
RL	Nucleic Acids Res. 15:3595-3606(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90355928; PubMed=2143802;			
RA	Chandra G., Gray J.G.;			
RT	"Cloning and expression of avidin in Escherichia coli.;"			
RL	Meth. Enzymol. 184:70-79(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=WHITE LEHORN; TISSUE=Oviduct;			
RX	MEDLINE=95394357; PubMed=7665080;			
RA	Wallen M.J., Laukkanen M.O., Kulomaa M.S.;			
RT	"Cloning and sequencing of the chicken egg-white avidin-encoding gene			
RT	and its relationship with the avidin-related genes Avrl-Avr5.;"			
RL	Gene 161:205-209(1995).			
RN	[4]			
RP	SEQUENCE OF 25-152.			
RX	MEDLINE=71107558; PubMed=5100763;			
RA	Delange R.J., Huang T.-S.;			
RT	"Egg white avidin. 3. Sequence of the 78-residue middle cyanogen			
RT	bromide peptide. Complete amino acid sequence of the protein			
RL	subunit.;"			
RL	J. Biol. Chem. 246:698-709(1971).			
RN	[5]			
RP	IMPORTANCE OF TYR IN BIOTIN-BINDING.			
RX	MEDLINE=90351377; PubMed=2386489;			
RA	Gitlin G., Bayer E.A., Wilchek M.;			
RT	"Studies on the biotin-binding sites of avidin and streptavidin.			
RT	Tyrosine residues are involved in the binding site.;"			
RL	Biochem. J. 269:527-530(1990).			
RN	[6]			
RP	BIOTIN-BINDING STUDIES.			
RX	MEDLINE=91378911; PubMed=1898347;			
RA	Hiller Y., Bayer E.A., Wilchek M.;			
RT	"Studies on the biotin-binding site of avidin. Minimized fragments			
RT	that bind biotin.;"			

ALIGNMENTS

34	68	8.0	1953	1	BIGA_SALTY	P25927 salmonella
35	67.5	8.0	489	1	RHAB_ECOLI	P32171 escherichia
36	67.5	8.0	565	1	YHXB_BACSU	P18159 bacillus su
37	67.5	8.0	1852	1	CCAS_CYPCA	P23216 cyprinus ca
38	67	7.9	482	1	K2C8_HUMAN	P03787 homo sapien
39	67	7.9	1033	1	TIRI_ECOLI	P10486 escherichia
40	66.5	7.9	210	1	RA10_YEAST	P06838 saccharomyc
41	66.5	7.9	306	1	CI81_HUMAN	O15165 homo sapien
42	66	7.8	369	1	YMG6_YEAST	Q04658 saccharomyc
43	66	7.8	417	1	PVR_CERAE	P32506 cercopitheci
44	66	7.8	911	1	CTDI_MOUSE	P30999 mus musculu
45	65.5	7.8	393	1	PGLR_PRUPE	P48979 prunus pers

RL Biochem. J. 278:573-585(1991).
 RN [7]
 RD X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93281699; PubMed=8506353;
 RA Livnah O., Bayer E.A., Wilchek M., Sussman J.L.;
 RT "Three-dimensional structures of avidin and the avidin-biotin
 complex.";
 RL PROC. Natl. Acad. Sci. U.S.A. 90:5076-5080(1993).
 RN [8]
 RD X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE=93294833; PubMed=8515446;
 RA Pugliese L., Coda A., Malcovati M., Bolognesi M.;
 RT "Three-dimensional structure of the tetragonal crystal form of
 egg-white avidin in its functional complex with biotin at 2.7-A
 resolution.";
 RL J. Mol. Biol. 231:698-710(1993).
 RN [9]
 RD X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=98430987; PubMed=9760187;
 RA Nardone E., Rosano C., Santambrogio P., Curnis F., Corti A., Magni F.,
 RA Siccardi A.G., Paganeli G., Losso R., Aprea B., Bolognesi M.,
 RA Sidoli A., Arosio P.;
 RT "Biochemical characterization and crystal structure of a recombinant
 hen avidin and its acidic mutant expressed in *Escherichia coli*.";
 RL Eur. J. Biochem. 256:453-460(1998).
 CC -!- FUNCTION: THE BIOLOGICAL FUNCTION OF AVIDIN IS NOT KNOWN. FORMS A
 CC STRONG NON-COVALENT SPECIFIC COMPLEX WITH BIOTIN (ONE MOLECULE OF
 CC BIOTIN PER SUBUNIT OF AVIDIN).
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN HEN OVIDUCT AND CONCENTRATED IN
 CC EGG WHITE (WHERE IT REPRESENT 0.05% OF THE TOTAL PROTEIN).
 CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
 CC -!- DATABASE: NAME=worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/AV.html".
 CC -!- DATABASE: NAME=prozyme technical fact sheet;
 CC WWW="http://www.prozyme.com/technical/av10data.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05343; CAA28954.1; -;
 CC EMBL; L27818; AAB59733.1; -;
 CC FIR; A03160; VICH.
 CC FIR; A27518; A27518.
 CC PIR; S11540; S11540.
 CC PDB; 2AVI; 15-JUL-93.
 CC PDB; 1AVD; 31-JAN-94.
 CC PDB; 1AVE; 31-JAN-94.
 CC PDB; 1RAV; 15-JUL-98.
 CC PDB; 2CAM; 15-JUL-98.
 CC InterPro; IPR000088; Avidin.
 CC Pfam; pf01382; Avidin; 1.
 CC PRINTS; PR00709; AVIDIN.
 CC ProDom; PD016055; Avidin; 1.
 CC PROSITE; PS00577; AVIDIN; 1.
 CC GlycoProtein; Signal; Biotin; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 152 AVIDIN.
 FT DISULFID 28 107 N-LINKED (GLCNAC...).
 FT CARBOHYD 41 41 BIOTIN (NON-COVALENT).
 FT BINDING 57 57 I -> T (IN APPROX. 50% OF THE CHAINS).
 FT VARIANT 58 58 G -> S (IN REF. 3).
 FT CONFLICT 22 22 E -> Q (IN REF. 2 AND 3).
 FT CONFLICT 77 77
 FT STRAND 32 36
 FT TURN 37 38
 FT STRAND 41 44
 FT TURN 49 50

FT STRAND 54 58
 FT STRAND 71 74
 FT STRAND 77 77
 FT TURN 83 84
 FT STRAND 87 93
 FT STRAND 100 109
 FT STRAND 115 124
 FT HELIX 130 135
 FT STRAND 137 146
 SQ SEQUENCE 152 AA; 16769 MW; 1D55A4491D5BFD5C CRC64;

 Query Match 80.6%; Score 681; DB 1; Length 152;
 Best Local Similarity 97.0%; Pred. No. 5.2e-61;
 Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 QY 30 PEIPARKGSLCKWTFNDLGSNNMTICAVNSRGSEFTGTYITAVTATSNKESPLHGTQNTI 89
 Db 21 PCLSARKCSLTKGKWNLDGSSNMTICAVNSRGSEFTGTYITAVTATSNKESPLHGTQNTI 80
 QY 90 NKRTQPTFGFTYNNKFSSESTTVFTGQCFIDRNGKEVLTMTLLRSVNDIGDDWKATRYG 149
 Db 81 NKRTQPTFGFTYNNKFSSESTTVFTGQCFIDRNGKEVLTMTLLRSVNDIGDDWKATRYG 140
 QY 150 INIFRLRTPQKE 161
 Db 141 INIFRLRTPQKE 152

 RESULT 2
 AVRA_CHICK STANDARD; PRT; 150 AA.
 ID AVRA_CHICK
 AC P56734;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Avidin-related protein 4/5 precursor.
 GN AVRA AND AVRS.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEHORN; TISSUE=Oviduct;
 RX MEDLINE=94170814; PubMed=8125122;
 RA Keinaenen R.A., Wallen M.J., Kristo P.A., Laukkanen M.O.,
 RA Toimela T.A., Helenius M.A., Kulomaa M.S.;
 RT "Molecular cloning and nucleotide sequence of chicken avidin-related
 genes 1-5.";
 RL Eur. J. Biochem. 220:615-621(1994).
 CC -!- MISCELLANEOUS: THE SEQUENCES OF THE CODING REGIONS OF GENES AVRA
 CC AND AVRS ARE IDENTICAL.
 CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z22883; -; NOT_ANNOTATED_CDS.
 CC HSSP; P02701; 1RAV.
 CC InterPro; IPR000088; Avidin.
 CC Pfam; PF01382; Avidin; 1.
 CC PRINTS; PR00709; AVIDIN.
 CC ProDom; PD016055; Avidin; 1.
 CC PROSITE; PS00577; AVIDIN; FALSE_NEG.
 KW Biotin; Signal; Multigene family.
 FT SIGNAL 1 24
 FT SIGNAL 1 24 POTENTIAL.

```
FT CHAIN 25 150 AVIDIN-RELATED PROTEIN 4/5.
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16644 MW; 9A6C6C6310EFEL3A CRC64;

Query Match 59.1%; Score 499; DB 1; Length 150;
Best Local Similarity 75.0%; Pred. No. 7.2e-43;
Matches 99; Conservative 8; Mismatches 23; Indels 2; Gaps 1;

QY 30 PEIPARKCSLTGKWTNDLGSNMNTIGAVNSRGEGFTGYITAVTATSNIEIKESPLHGQTNTI 89
Db 21 PSLARKCSLTGKWTNNLGSIMTIRAVNSRGEGFTGYLTAVADNPGNIKLSPLLGIOH-- 78
QY 90 NKRTQPTFGFTVNNKFSSESTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
Db 79 KRASQPTFGFTVHWNFSESTVFTGQCFIDRNGKEVLKTMWLLRSSVNDISYDWKATRVG 138
QY 150 INIFTRLTQKE 161
Db 139 YNNFTRLTQVEE 150

RESULT 3
AVR3_CHICK
ID AVR3_CHICK STANDARD; PRT; 150 AA.
AC P56733;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Avidin-related protein 3 precursor.
GN AVR3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE=Oviduct;
RX MEDLINE=94170814; PubMed=8125122;
RA Keinonen R.A., Wallen M.J., Kristo P.A., Laukkanen M.O.,
RA Toimela T.A., Helenius M.A., Kulomaa M.S.;
RT Molecular cloning and nucleotide sequence of chicken avidin-related
genes 1-5.*;
RL Eur. J. Biochem. 220:615-621(1994).
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; 221612; -; NOT_ANNOTATED_CDS.
DR EMBL; 221536; -; NOT_ANNOTATED_CDS.
DR HSP; P02701; 1RAV.
DR InterPro; IPR000088; Avidin.
DR Pfam; PF01382; Avidin; 1.
DR PRINTS; PR00709; AVIDIN.
DR PRODOM; PD016055; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.
KW Biotin; Signal; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 150 AVIDIN-RELATED PROTEIN 3.
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16657 MW; 84DC03926A6BE21C CRC64;
```

```
Query Match 55.5%; Score 469; DB 1; Length 150;
Best Local Similarity 69.7%; Pred. No. 7e-40;
Matches 92; Conservative 11; Mismatches 27; Indels 2; Gaps 1;

QY 30 PEIPARKCSLTGKWTNDLGSNMNTIGAVNSRGEGFTGYITAVTATSNIEIKESPLHGQTNTI 89
Db 21 PSLARKCSLTGKWTNNLGSIMTIRAVNSRGEGFTGYLTAVADNPGNIKLSPLLGIOH-- 78
QY 90 NKRTQPTFGFTVNNKFSSESTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
Db 79 KRASQPTFGFTVHWNFSESTVFTGQCFIDRNGKEVLKTMWLLRSSVNDISDWKATRVG 138
QY 150 INIFTRLTQKE 161
Db 139 YNNFTRLTQVEE 150

RESULT 4
AVR7_CHICK
ID AVR7_CHICK STANDARD; PRT; 150 AA.
AC P56736; Q9W6V4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Avidin-related protein 7 precursor.
GN AVR7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RHODE ISLAND;
RA Ahlroth M.K., Kola E.H., Ewald D., Masabanda J., Sazanov A., Fries R.,
RA Kulomaa M.S.;
RT "Characterization and chromosomal localization of the chicken avidin
gene family.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AJ237659; CAB39894.1; -
DR HSP; P02701; 1RAV.
DR InterPro; IPR000088; Avidin.
DR Pfam; PF01382; Avidin; 1.
DR PRINTS; PR00709; AVIDIN.
DR PRODOM; PD016055; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.
KW Biotin; Signal; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 150 AVIDIN-RELATED PROTEIN 7.
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16543 MW; 30F335611F65923D CRC64;

Query Match 55.3%; Score 467; DB 1; Length 150;
Best Local Similarity 68.9%; Pred. No. 1.1e-39;
Matches 91; Conservative 11; Mismatches 28; Indels 2; Gaps 1;
```



```

Matches 33; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY 1 MESKFAHIIIVFFLLATPFETLLARKESDGPEI-----PARKSLTGKWTNDLG 48
   |||||
Db 1 MESKFAHIIIVFFLLATPFETLLARKESDGPEVIELQKEFCNGKQRPDELIG 52
   |||||

RESULT 9
ICIL_SOLITU
ID TOTL_SOLITU STANDARD; PRT; 107 AA.
AC Q00783;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Proteinase inhibitor I precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RD SEQUENCE FROM N.A.
RC STRAIN=CV, RUSSET BURBANK;
RA Lee J.S.; Park J.S.;
RT "Nucleotide sequence of a protease inhibitor I gene in potato.";
RL Korean J. Bot. 32:67-78(1989).
CC -!- SIMILARITY: BELONGS TO THE POTATO CHYMOTRYPSIN INHIBITOR I FAMILY
CC OF SERINE PROTEASE INHIBITOR.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z12611; CAA79259.1; -.
DR HSSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHIBTR.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT PROPEP 24 36 BY SIMILARITY.
FT CHAIN 37 107 PROTEINASE INHIBITOR I.
FT ACT_SITE 83 84 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 107 AA; 12063 MW; 52F39C3C913CC052 CRC64;

Query Match 16.9%; Score 143; DB 1; Length 107;
Best Local Similarity 87.5%; Pred. No. 1.5e-07;
Matches 28; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MESKFAHIIIVFFLLATPFETLLARKESDGPEI 32
   |||||
Db 1 MELKFAHIIIVFFLLATSFETLLARKESDGPEV 32
   |||||

RESULT 10
ICIL_LYCES
ID ICIL_LYCES STANDARD; PRT; 111 AA.
AC P05118;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Wound-induced proteinase inhibitor I precursor.
OS Lycopersicon esculentum (Tomato); and
OS Lycopersicon peruvianum (Peruvian Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

```

OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081, 4082;
RN [1]
RD SEQUENCE FROM N.A.
RC SPECIES=L.esculentum;
RX MEDLINE=87016930; PubMed=3463966;
RA Lee J.S., BROWN W.E., Graham J.S., Pearce G., Fox E.A., Dreher T.W.,
RA Ahern K.G., Pearson G.D., Ryan C.A.;
RT "Molecular characterization and phylogenetic studies of a wound-
RL inducible proteinase inhibitor I gene in Lycopersicon species.";
RN [2]
RD SEQUENCE FROM N.A.
RC SPECIES=L.esculentum;
RX MEDLINE=85207657; PubMed=2987227;
RA Graham J.S., Pearce G., Merryweather J., Titani K., Ericsson L.,
RA Ryan C.A.;
RT "Wound-induced proteinase inhibitors from tomato leaves. I. The cDNA-
RT deduced primary structure of pre-inhibitor I and its
RT post-translational processing.";
RL J. Biol. Chem. 260:6555-6560(1985).
RN [3]
RD SEQUENCE FROM N.A.
RC SPECIES=L.peruvianum;
RX MEDLINE=91170261; PubMed=2005119;
RA Wingate V.P.M., Ryan C.A.;
RT "A novel fruit-expressed trypsin inhibitor I gene from a wild species
RT of tomato.";
RL J. Biol. Chem. 266:5814-5818(1991).
CC -!- MISCELLANEOUS: MECHANICAL DAMAGE (I.E., INSECT CHEWING) TO THIS
CC PLANT RESULTS IN THE SYSTEMIC RELEASE OF A FACTOR FROM THE WOUND
CC SITE. WITHIN THE LEAVES IT INDUCES THE CYTOPLASMIC SYNTHESIS OF
CC PROTEINASE INHIBITORS I AND II.
CC -!- SIMILARITY: BELONGS TO THE POTATO CHYMOTRYPSIN INHIBITOR I FAMILY
CC OF SERINE PROTEASE INHIBITOR.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K03290; AAA34199.1; -.
DR EMBL; M13938; AAA34200.1; -.
DR EMBL; M59427; AAA34198.1; ALT_SEQ.
DR PIR; A24048; A24048.
DR PIR; A25046; A25046.
DR PIR; A39547; A39547.
DR HSSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHIBTR.
DR PRODOM; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 23
FT PROPEP 24 36
FT CHAIN 37 111 WOUND-INDUCED PROTEINASE INHIBITOR I.
FT ACT_SITE 87 88 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 111 AA; 12585 MW; 124BAC3400E9F0D6 CRC64;

Query Match 16.9%; Score 143; DB 1; Length 111;
Best Local Similarity 87.5%; Pred. No. 1.5e-07;
Matches 28; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MESKFAHIIIVFFLLATPFETLLARKESDGPEI 32
   |||||
Db 1 MESKFAHIIIVFFLLATSFETLLARKESDGPEV 32
   |||||

```

```

RESULT 11
SAVL_STRLV          STANDARD;          PRT;    183 AA.
ID   SAVL_STRLV
AC   Q53532;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DE   Streptavidin VI precursor (SA V1).
OS   Streptomyces violaceus (Streptomyces venezuelae).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1936;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95359204; PubMed=7632734;
RA   Bayer E.A., Kulik T., Adar R., Wilchek M.;
RT   "Close similarity among streptavidin-like, biotin-binding proteins
RL   from Streptomyces.";
RL   Biochim. Biophys. Acta 1263:60-66(1995).
CC   FUNCTION: THE BIOLOGICAL FUNCTION OF STREPTAVIDIN IS NOT KNOWN.
CC   FORMS A STRONG NON-COVALENT SPECIFIC COMPLEX WITH BIOTIN (ONE
CC   MOLECULE OF BIOTIN PER SUBUNIT OF STREPTAVIDIN).
CC   SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   SUBCELLULAR LOCATION: Secreted.
CC   SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC   THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
BR   EMBL; S78777; AAB35015.1; -.
DR   HSSP; P22629; LSRI.
DR   InterPro; IPR000088; Avidin.
DR   Pfam; PF01382; Avidin; 1.
DR   ProDom; PD016055; Avidin; 1.
DR   PROSITE; PS00577; AVIDIN; 1.
KW   Signal; Biotin.
FT   SIGNAL          1      24      BY SIMILARITY.
FT   CHAIN           25     183     STREPTAVIDIN V1.
FT   BINDING         67     67     INVOLVED IN BIOTIN BINDING (BY
FT                                     SIMILARITY).
FT   BINDING         78     78     INVOLVED IN BIOTIN BINDING (BY
FT                                     SIMILARITY).
FT   BINDING         116    116     INVOLVED IN BIOTIN BINDING (BY
FT                                     SIMILARITY).
FT   BINDING         132    132     INVOLVED IN BIOTIN BINDING (BY
FT                                     SIMILARITY).
FT   BINDING         144    144     INVOLVED IN BIOTIN BINDING (BY
FT                                     SIMILARITY).
FT   SEQUENCE        183 AA; 18864 MW; 4652D8B814847DE0 CRC64;
SQ

```

Query Match 16.9%; Score 142.5; DB 1; Length 183;

Best Local Similarity 32.3%; Pred. No. 3.2e-07;

Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

```

QY 31 ETPARKSLTGKWNLDGSMRTGAVNSRGFTGYITAV-TATSNKEIKESPLHGTONTI 89
DB 33 QVSAEAGITGYWYNOLGSTFFIVA-GAGCALTGTVESAVGNRESYVLTRGYDSAPATD 91
QY 90 NKRTPTFGFTYWNK-----FSESTTVFTGCFTDRNGKEY-LKTMILLRSSVNDIGDWWK 144
DB 92 GSST--ALGWTVAWKNYRNAHSATTSQYV---GGTEARINTQWLLTSGTTE-ANAWK 145
QY 145 AFRVGINIFTRLR 157
DB 146 STLVGHDFTTKVK 158

```

```

RESULT 12
SAV_STRAV          STANDARD;          PRT;    183 AA.
ID   SAV_STRAV
AC   P22629;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DE   Streptavidin precursor.
OS   Streptomyces avidinii.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1895;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 25-64.
RX   MEDLINE=86148514; PubMed=3951999;
RA   Argarana C.E., Kuntz I.D., Birken S., Axel R., Cantor C.R.;
RT   "Molecular cloning and nucleotide sequence of the streptavidin gene.";
RL   Nucleic Acids Res. 14:1871-1882(1986).
RN   [2]
RP   IMPORTANCE OF TRP IN BIOTIN-BINDING.
RX   MEDLINE=89134083; PubMed=3223904;
RA   Gitlin G., Bayer E.A., Wilchek M.;
RT   "Studies on the biotin-binding site of streptavidin. Tryptophan
RL   residues involved in the active site.";
RL   Biochem. J. 256:279-282(1988).
RN   [3]
RP   IMPORTANCE OF TYR IN BIOTIN-BINDING.
RX   MEDLINE=90351377; PubMed=2386489;
RA   Gitlin G., Bayer E.A., Wilchek M.;
RT   "Studies on the biotin-binding sites of avidin and streptavidin.
RL   Tyrosine residues are involved in the binding site.";
RL   Biochem. J. 269:527-530(1990).
RN   [4]
RP   CELL-BINDING.
RX   MEDLINE=90358825; PubMed=2390089;
RA   Alon R., Bayer E.A., Wilchek M.;
RT   "Streptavidin contains an RYD sequence which mimics the RGD receptor
RL   domain of fibronectin.";
RL   Biochem. Biophys. Res. Commun. 170:1236-1241(1990).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 40-157.
RX   MEDLINE=89100243; PubMed=2911722;
RA   Weber P.C., Ohlendorf D.H., Wendoloski J.J., Salem F.R.;
RT   "Structural origins of high-affinity biotin binding to streptavidin.";
RL   Science 243:85-88(1989).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 40-157.
RX   MEDLINE=97337436; PubMed=9194176;
RA   Freitag S., le Trong I., Klumb L., Stayton P.S., Stenkamp R.E.;
RT   "Structural studies of the streptavidin binding loop.";
RL   Protein Sci. 6:1157-1166(1997).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 37-157.
RX   MEDLINE=97294734; PubMed=9148939;
RA   Katz B.A., Cass R.T.;
RT   "In crystals of complexes of streptavidin with peptide ligands
RL   containing the HPQ sequence the pKa of the peptide histidine is less
RL   than 3.0.";
RL   J. Biol. Chem. 272:13220-13228(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 37-157.
RX   MEDLINE=98070615; PubMed=9405158;
RA   Katz B.A.;
RT   "Binding of biotin to streptavidin stabilizes intersubunit salt
RL   bridges between Asp61 and His87 at low pH.";
RL   J. Mol. Biol. 274:776-800(1997).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 37-157 OF MUTANTS.
RX   MEDLINE=98300340; PubMed=9636711;
RA   Freitag S., le Trong I., Chilkoti A., Klumb L.A., Stayton P.S.,
RA   Stenkamp R.E.;
RT   "Structural studies of binding site tryptophan mutants in the high-
RL   affinity streptavidin-biotin complex.";

```



```

FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT DISULFID 446 457 BY SIMILARITY.
FT DISULFID 451 466 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 489 504 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 522 533 BY SIMILARITY.
FT DISULFID 527 542 BY SIMILARITY.
FT DISULFID 544 553 BY SIMILARITY.
FT DISULFID 560 571 BY SIMILARITY.
FT DISULFID 565 580 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 603 618 BY SIMILARITY.
FT DISULFID 620 629 BY SIMILARITY.
FT DISULFID 636 647 BY SIMILARITY.
FT DISULFID 641 656 BY SIMILARITY.
FT DISULFID 658 667 BY SIMILARITY.
FT DISULFID 674 685 BY SIMILARITY.
FT DISULFID 679 694 BY SIMILARITY.
FT DISULFID 696 705 BY SIMILARITY.
FT DISULFID 712 723 BY SIMILARITY.
FT DISULFID 717 732 BY SIMILARITY.
FT DISULFID 734 743 BY SIMILARITY.
FT DISULFID 750 761 BY SIMILARITY.
FT DISULFID 755 770 BY SIMILARITY.
FT DISULFID 772 781 BY SIMILARITY.
FT DISULFID 788 799 BY SIMILARITY.
FT DISULFID 793 808 BY SIMILARITY.
FT DISULFID 810 819 BY SIMILARITY.
FT DISULFID 826 837 BY SIMILARITY.
FT DISULFID 831 846 BY SIMILARITY.
FT DISULFID 848 857 BY SIMILARITY.
FT DISULFID 864 875 BY SIMILARITY.
FT DISULFID 869 884 BY SIMILARITY.
FT DISULFID 886 895 BY SIMILARITY.
FT DISULFID 902 913 BY SIMILARITY.
FT DISULFID 907 922 BY SIMILARITY.
FT DISULFID 924 933 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 851 851 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPHIC 477 780 MISSING (IN ISOFORM IB).
FT CONFLICT 279 279 L -> S (IN REF. 2).
SQ SEQUENCE 1064 AA; 112072 MW; 2E569CA012ED6D09 CRC64;

Query Match 16.6%; Score 140; DB 1; Length 1064;
Best Local Similarity 33.6%; Pred. NO. 5.2e-06;
Matches 42; Conservative 16; Mismatches 51; Indels 16; Gaps 4;

QY 37 CSLTCKWTFNDLGSNMWTCICAVNSRCGFTCTYITAVTATSNKEIKESPLHGTQNTI-----N 90
DB 939 CDLEGMHNECNDQYTI-TKTSQNMGLGDYMI-----NTERALGYAAFTVYVGYASN 989
QY 91 KRTPQTPGVNRFSESTVFTGQCFIDRNGKEVLRKPMWILRSSVNDIGDDWKATRVG 150
DB 990 NYDFPSGFTVVRDNGOSTTSWTGCOHL-CDGEELYLTWTWNTNMVSTCODIKSNMVGQ 1048
QY 151 NIFTR 155
DB 1049 DKWTR 1053

RESULT 14
SAV2_STRVL
-ID SAV2_STRVL STANDARD; PRT; 183 AA.

```

```

AC Q55533;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Streptavidin V2 precursor (SA V2)
OS Streptomyces violaceus (Streptomyces venezuelae).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID-19336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95359204; PubMed-7632734;
RA Bayer E.A.; Kulik R.; Adair R.; Wilchek M.;
RT "Close similarity among streptavidin-like, biotin-binding proteins
RL Biochim. Biophys. Acta 1263:60-66(1995).
CC -!- FUNCTION: THE BIOLOGICAL FUNCTION OF STREPTAVIDIN IS NOT KNOWN.
CC FORMS A STRONG NON-COVALENT SPECIFIC COMPLEX WITH BIOTIN (ONE
CC MOLECULE OF BIOTIN PER SUBUNIT OF STREPTAVIDIN).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S78782; AAB35016.1; .
DR HSP; P22629; IPTS.
DR InterPro; IPR000088; Avidin.
DR Pfam; pf01382; Avidin; 1.
DR PRODOM; PD016055; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.
KW signal; biotin.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 183 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 67 67 SIMILARITY).
FT BINDING 78 78 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 116 116 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 132 132 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 144 144 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 183 AA; 18833 MW; FEAFFDFEA4ECCA CRC64;
SQ SEQUENCE 183 AA; 18833 MW; FEAFFDFEA4ECCA CRC64;

Query Match 16.5%; Score 139.5; DB 1; Length 183;
Best Local Similarity 33.6%; Pred. NO. 6.4e-07;
Matches 43; Conservative 17; Mismatches 55; Indels 13; Gaps 7;

QY 36 KCSLTCKWTFNDLGSNMWTCICAVNSRCGFTCTYITAV-TATSNKEIKESPLHGTQNTINKRTO 94
DB 30 EAGITGTWYINQUGSTFTIYTA-NADGSLTGTYESAVNGESRYLVTGRYDSAFATDGGT- 95
QY 95 PREGFTVMRK----PSESTVFTGQCFIDRNGKEV-LRPMWILRSSVNDIGDDWKATRVG 149
DB 96 -ALGWTVAWKNNYRNAHSATTSWGGYVA---GSEARINTQWLLTSGTT-AANAWKSTLVG 150
QY 150 NIFTR 157
DB 151 HDFTKVK 158

RESULT 15
ICIL_LYCPE
ID ICIL_LYCPE STANDARD; PRT; 111 AA.

```

```

AC P16231;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update).
DE Wound-induced proteinase inhibitor I precursor.
OS Lycopersicon peruvianum (Peruvian tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90036835; PubMed=2808345;
RA Wingate V.P.M., Broadway R.M., Ryan C.A.;
RT "Isolation and characterization of a novel, developmentally regulated
RT proteinase inhibitor I protein and cDNA from the fruit of a wild
RT species of tomato.";
RL J. Biol. Chem. 264:17734-17738(1989).
CC -!- SIMILARITY: BELONGS TO THE POTATO CHYMOTRYPSIN INHIBITOR I FAMILY
CC OF SERINE PROTEASE INHIBITOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05094; AAA34180.1; -.
DR PIR; A34359; A34359.
DR HSP; P19873; LHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR PRINTS; PR00292; POTATOINHBT.
DR ProDom; PD002604; Potato_inhibit; 1.
DR PROSITE; PS00285; POTATO_INHIBITOR; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 23
FT PROPEP 24 36
FT CHAIN 37 111
FT ACT_SITE 87 88 WOUND-INDUCED PROTEINASE INHIBITOR I.
SQ SEQUENCE 111 AA; 12610 MW; 3777F7C63231C265 CRC64;

```

```

Query Match      16.1%; Score 136; DB 1; Length 111;
Best Local Similarity 81.2%; Pred. No. 7.6e-07;
Matches 26; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MESKFAHIIIVFFLLATPFETLLARKESDGPFI 32
   |||||||:||||| ||||:|||||
DB 1 MEAKFAHIIIFLLAFSFETLLARKESDGPV 32

```

Search completed: September 17, 2002, 15:43:39
Job time: 125 sec

A:Reference number: A51622; PDB:1AVD
 A:Contents: annotation; X-ray crystallography, 2.7 angstroms, with biotin, residues 27-3
 R:Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M.
 submitted to the Brookhaven Protein Data Bank, March 1993
 A:Reference number: A51623; PDB:1AVE
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms, without biotin, residues 2
 R:Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M.
 J. Mol. Biol. 231, 698-710, 1993
 A:Title: Three-dimensional structure of the tetragonal crystal form of egg-white avidin
 A:Reference number: A54974; MUID:93294833
 A:Contents: annotation; X-ray crystallography, 2.7 angstroms
 C:Genetics:
 A:Introns: 27/3; 96/1; 136/2
 C:Superfamily: avidin
 C:Keywords: glycoprotein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-152/Product: avidin #status experimental <MAT>
 F:28-107/Disulfide bonds: #status experimental
 F:41/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 80.9%; Score 684; DB 1; Length 152;
 Best Local Similarity 97.7%; Pred. No. 7.3e-60;
 Matches 129; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 30 PEIPARKCSLTGKWTNDLGSNMITGAVNSRGEFTGYITAVTATSNKEIKESPLHGQTNTI 89
 I : |||||
 DB 21 PCLSARKCSLTGKWTNDLGSNMITGAVNSRGEFTGYITAVTATSNKEIKESPLHGQTNTI 80
 I : |||||

QY 90 NKRTQPTFGFTVNNKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
 I : |||||
 DB 81 NKRPQPTFGFTVNNKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 140
 I : |||||

QY 150 INIFTRLRTOKE 161
 I : |||||
 DB 141 INIFTRLRTOKE 152
 I : |||||

RESULT 2
 S42204
 A:Title: Avidin-related protein 4/5 precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
 C:Accession: S42204; S42205
 R:Keinaenen, R.A.; Wallen, M.J.; Kristo, P.A.; Laukkanen, M.O.; Toimela, T.A.; Helenius,
 Eur. J. Biochem. 220, 615-621, 1994
 A:Title: Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5.
 A:Reference number: S42201; MUID:94170814
 A:Accession: S42204
 A:Molecule type: DNA
 A:Residues: 1-150 <KEI>
 A:Cross-references: EMBL:Z22883; NID:q311811
 A:Experimental source: strain White Leghorn; tissue oviduct
 A:Genetics: CH1
 A:Accession: S42205
 A:Molecule type: DNA
 A:Residues: 1-150 <KEW>
 A:Cross-references: EMBL:Z22882; NID:q311812
 A:Experimental source: strain White Leghorn; tissue oviduct
 A:Genetics: CH2
 A:Gene: avr4
 A:Introns: 27/3; 96/1; 136/2
 C:Genetics: <CH2>
 A:Gene: avr5
 A:Introns: 27/3; 96/1; 136/2
 C:Superfamily: avidin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-150/Product: avidin-related protein 4/5 #status predicted <MAT>
 F:28-105/Disulfide bonds: #status predicted

Query Match 59.1%; Score 499; DB 2; Length 150;

Best Local Similarity 75.0%; Pred. NO. 1e-41;
 Matches 99; Conservative 8; Mismatches 23; Indels 2; Gaps 1;

QY 30 PEIPARKCSLTGKWTNDLGSNMITGAVNSRGEFTGYITAVTATSNKEIKESPLHGQTNTI 89
 I : |||||
 DB 21 PCLSARKCSLTGKWTNNLGSIMTIRAVNSRGEFTGYITAVADNPGNIKLSPLLGIOH-- 78
 I : |||||

QY 90 NKRTQPTFGFTVNNKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
 I : |||||
 DB 79 KRASQPTFGFTVNNKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDISYDOKATRVG 138
 I : |||||

QY 150 INIFTRLRTOKE 161
 I : |||||
 DB 139 YNNFTRLCTVEE 150
 I : |||||

RESULT 3
 S42203
 A:Title: Avidin-related protein 3 precursor - chicken
 N:Alternate names: avr3 protein
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
 C:Accession: S42203; S39800
 R:Keinaenen, R.A.; Wallen, M.J.; Kristo, P.A.; Laukkanen, M.O.; Toimela, T.A.; Heleni
 Eur. J. Biochem. 220, 615-621, 1994
 A:Title: Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-
 A:Reference number: S42201; MUID:94170814
 A:Accession: S42203
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-150 <KEI>
 A:Cross-references: EMBL:Z21612; NID:g65432
 R:Kunna, T.A.; Wallen, M.J.; Kulomaa, M.S.
 Biochim. Biophys. Acta 1216, 441-445, 1993
 A:Title: Induction of chicken avidin and related mRNAs after bacterial infection.
 A:Reference number: S39799; MUID:94092737
 A:Accession: S39800
 A:Status: preliminary; translation not shown
 A:Molecule type: mRNA
 A:Residues: 71-150 <KUN>
 A:Cross-references: EMBL:Z21536; NID:g65429
 C:Genetics:
 A:Gene: avr3
 A:Introns: 27/3; 96/1; 136/2
 C:Superfamily: avidin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-150/Product: avidin-related protein 3 #status predicted <MAT>
 F:28-105/Disulfide bonds: #status predicted

Query Match 55.5%; Score 469; DB 2; Length 150;
 Best Local Similarity 69.7%; Pred. No. 9.1e-39;
 Matches 92; Conservative 11; Mismatches 27; Indels 2; Gaps 1;

QY 30 PEIPARKCSLTGKWTNDLGSNMITGAVNSRGEFTGYITAVTATSNKEIKESPLHGQTNTI 89
 I : |||||
 DB 21 PCLSARKCSLTGKWTNNLGSIMTIRAVNSRGEFTGYITAVADNPGNIKLSPLLGIOH-- 78
 I : |||||

QY 90 NKRTQPTFGFTVNNKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
 I : |||||
 DB 79 KRACQPTFGFTVNNKFSSTVFTGQCIFDRSGKEVLKTKWLORLAVDDISDDWKATRVG 138
 I : |||||

QY 150 INIFTRLRTOKE 161
 I : |||||
 DB 139 YNNFTRLCTVEE 150
 I : |||||

RESULT 4
 S42201
 A:Title: Avidin-related protein 1 precursor - chicken
 N:Alternate names: avr1
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999

A:Residues: 1-107 <BEU>
A:Cross-references: EMBL:X67675; NID:g21527; PIDN:CAA47907.1; PID:g21528; EMBL:
A:Experimental source: cultivar Ilam Hardy

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:39:39 ; Search time 12.95 seconds
(without alignments)
303.670 Million cell updates/sec

Title: US-09-743-690-7

Perfect score: 845

Sequence: 1 MESKFAHIIFFLLATPFET.....DWKATRGVNIIFRLTQKE 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	684	80.9	152	4 US-08-831-399-4	Sequence 4, Appli
2	142.5	16.9	159	3 US-08-628-540-1	Sequence 1, Appli
3	142.5	16.9	159	3 US-08-948-097-17	Sequence 17, Appl
4	142.5	16.9	159	4 US-08-941-100-1	Sequence 1, Appli
5	142.5	16.9	159	4 US-08-941-100-5	Sequence 5, Appli
6	142.5	16.9	183	4 US-08-831-399-2	Sequence 2, Appli
7	142.5	16.9	183	5 PCT-US93-05240-14	Sequence 14, Appl
8	142.5	16.9	183	6 5168049-5	Patent No. 5168049
9	142.5	16.9	415	2 US-08-491-988-7	Sequence 7, Appli
10	142.5	16.9	435	2 US-08-491-988-5	Sequence 5, Appli
11	141	16.7	402	2 US-08-491-988-9	Sequence 9, Appli
12	135.5	16.0	128	1 US-08-211-833-2	Sequence 2, Appli
13	135.5	16.0	128	1 US-08-434-718-2	Sequence 2, Appli
14	135.5	16.0	128	4 US-08-831-399-16	Sequence 16, Appl
15	115	13.6	21	1 US-08-554-586-2	Sequence 2, Appli
16	115	13.6	21	1 US-08-554-586-3	Sequence 3, Appli
17	67	7.9	462	2 US-08-752-307B-7	Sequence 7, Appli
18	67	7.9	465	2 US-08-752-307B-5	Sequence 5, Appli
19	66.5	7.9	997	1 US-08-232-540-1	Sequence 1, Appli
20	66.5	7.9	997	1 US-08-428-949A-1	Sequence 1, Appli
21	66.5	7.9	997	1 US-08-428-948A-1	Sequence 1, Appli
22	66.5	7.9	997	2 US-08-428-946-1	Sequence 1, Appli
23	66.5	7.9	997	5 PCT-US95-04656-1	Sequence 1, Appli
24	66.5	7.9	998	1 US-08-233-008A-6	Sequence 6, Appli
25	66.5	7.9	1021	1 US-08-233-008A-2	Sequence 2, Appli
26	66	7.8	682	4 US-08-982-785A-9	Sequence 9, Appli
27	66	7.8	703	4 US-08-910-925-4	Sequence 4, Appli

28	65.5	7.8	1257	1	US-08-340-428B-49	Sequence 49, Appli
29	64.5	7.6	325	1	US-08-292-549-2	Sequence 2, Appli
30	64.5	7.6	325	4	US-09-042-785A-9	Sequence 9, Appli
31	64.5	7.6	325	5	PCT-US91-02207-2	Sequence 2, Appli
32	64.5	7.6	674	1	US-08-317-522A-3	Sequence 3, Appli
33	64.5	7.6	674	1	US-08-439-818A-3	Sequence 3, Appli
34	64.5	7.6	674	2	US-08-751-965-3	Sequence 3, Appli
35	64.5	7.6	674	2	US-08-738-975-3	Sequence 3, Appli
36	64.5	7.6	674	2	US-08-728-626-3	Sequence 3, Appli
37	64.5	7.6	674	3	US-08-808-599A-3	Sequence 3, Appli
38	64.5	7.6	749	1	US-08-317-522A-2	Sequence 2, Appli
39	64.5	7.6	749	1	US-08-439-818A-2	Sequence 2, Appli
40	64.5	7.6	749	2	US-08-751-965-2	Sequence 2, Appli
41	64.5	7.6	749	2	US-08-738-975-2	Sequence 2, Appli
42	64.5	7.6	749	2	US-08-728-626-2	Sequence 2, Appli
43	64.5	7.6	749	3	US-08-808-599A-2	Sequence 2, Appli
44	64.5	7.6	2476	2	US-08-276-967-2	Sequence 2, Appli
45	64	7.6	630	1	US-08-487-890A-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-08-831-399-4
; Sequence 4, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Hans; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; TITLE OF INVENTION: Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 638-3884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-831-399-4

Query Match 80.9%; Score 684; DB 4; Length 152;
Best Local Similarity 97.7%; Pred. No. 2.2e-73;
Matches 129; Conservative 1; Mismatches 2; Indels 0; Gaps 0;


```
;
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)
; OTHER INFORMATION: The residue in this position can be any amino
; OTHER INFORMATION: acid.
US-08-941-100-1

Query Match          16.9%; Score 142.5; DB 4; Length 159;
Best Local Similarity 32.3%; Pred. No. 3.3e-09;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

Qy 31 EIPARKSLTGKWTNDLGSNMITGAVNSRGETGTITAV-TATSNKESPLHGQTNTI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 QVSAEAGITGTWYNLGSSTFIVTA-GADGALTGTYESAVGNAESRYVLTGRYDSAPATD 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 90 NKRTQPTFGFTVNWK----FSESTVFTGQCFFDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GSGT--ALGWTVAWKNNRYNAHSATWSGQYV---GGAEARINTQWLLTSGTTE-ANAWK 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 145 ATRVGINIFTRLR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 STLVGHDFTFKVK 134

RESULT 5
US-08-941-100-5
; Sequence 5, Application US/08941100B
; Patent No. 6207390
; GENERAL INFORMATION:
; APPLICANT: Cantor, Charles R.
; TITLE OF INVENTION: Reduced Affinity Streptavidin
; FILE REFERENCE: BU-03165
; CURRENT APPLICATION NUMBER: US/08/941,100B
; PRIOR FILING DATE: 1997-10-03
; PRIOR APPLICATION NUMBER: 08/469,353
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/420,010
; PRIOR FILING DATE: 1995-04-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Streptomyces avidinii
US-08-941-100-5

Query Match          16.9%; Score 142.5; DB 4; Length 159;
Best Local Similarity 32.3%; Pred. No. 3.3e-09;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

Qy 31 EIPARKSLTGKWTNDLGSNMITGAVNSRGETGTITAV-TATSNKESPLHGQTNTI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 QVSAEAGITGTWYNLGSSTFIVTA-GADGALTGTYESAVGNAESRYVLTGRYDSAPATD 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 90 NKRTQPTFGFTVNWK----FSESTVFTGQCFFDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 GSGT--ALGWTVAWKNNRYNAHSATWSGQYV---GGAEARINTQWLLTSGTTE-ANAWK 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 145 ATRVGINIFTRLR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 STLVGHDFTFKVK 134

RESULT 6
US-08-831-399-2
; Sequence 2, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
```

```
;
; TITLE OF INVENTION: Recombinant Inactive Core
; TITLE OF INVENTION: Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-831-399-2

Query Match          16.9%; Score 142.5; DB 4; Length 183;
Best Local Similarity 32.3%; Pred. No. 4e-09;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

Qy 31 EIPARKSLTGKWTNDLGSNMITGAVNSRGETGTITAV-TATSNKESPLHGQTNTI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 33 QVSAEAGITGTWYNLGSSTFIVTA-GADGALTGTYESAVGNAESRYVLTGRYDSAPATD 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 90 NKRTQPTFGFTVNWK----FSESTVFTGQCFFDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 GSGT--ALGWTVAWKNNRYNAHSATWSGQYV---GGAEARINTQWLLTSGTTE-ANAWK 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 145 ATRVGINIFTRLR 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 STLVGHDFTFKVK 158

RESULT 7
PCT-US93-05240-14
; Sequence 14, Application PC/TUS9305240
; GENERAL INFORMATION:
; APPLICANT: NAGARAJAN, VASANTHA
; TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN FROM BACILLUS
; TITLE OF INVENTION: SUBTILLIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DU PONT COMPANY
; STREET: BARLEY MILL PLAZA 36
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19880-0036
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/Ms-Dos
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/us93/05240
; FILING DATE: 19930527
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GEIGER, KATHLEEN W
; REFERENCE/DOCKET NUMBER: CK 9029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-692-2118
; TELEFAX: 302-692-7949
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-05240-14

      Query Match          16.9%; Score 142.5; DB 5; Length 183;
Best Local Similarity    32.3%; Pred. NO. 4e-09;
Matches                  43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

QY   31 EIPARKSLTGKWTNDLGSNMTIGAVNSRGEFTGTITAV-TATSNFEIKESPLHGTONTI 89
Db   33 QVSAAEAGITGTYNQLGSTFIYTA-GADGALTGTYESAVGNAESRYVLTRGYDSAPATD 91
QY   90 NKRTOPTFGFTVNWK-----PSESTVTGTGCFTDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
Db   92 GSGT--ALGWTVAKNNYRNAHSATTWSGOYV---GGAEARINTQWLLTSGTTT-ANAWK 145
QY   145 ATRVGINIFRLR 157
Db   146 STLGHDTFKVK 158

RESULT      8
5168049-5
; Patent No. 5168049
; APPLICANT: MEADE, HARRY M.; GARWIN, JEFFREY L.
; TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN-LIKE
; POLYPEPTIDES
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/185,329
; FILING DATE: 21-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,873
; FILING DATE: 02-OCT-1984
; SEQ ID NO: 5
; LENGTH: 183
5168049-5

      Query Match          16.9%; Score 142.5; DB 6; Length 183;
Best Local Similarity    32.3%; Pred. NO. 4e-09;
Matches                  43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

QY   31 EIPARKSLTGKWTNDLGSNMTIGAVNSRGEFTGTITAV-TATSNFEIKESPLHGTONTI 89
Db   33 QVSAAEAGITGTYNQLGSTFIYTA-GADGALTGTYESAVGNAESRYVLTRGYDSAPATD 91
QY   90 NKRTOPTFGFTVNWK-----PSESTVTGTGCFTDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
Db   92 GSGT--ALGWTVAKNNYRNAHSATTWSGOYV---GGAEARINTQWLLTSGTTT-ANAWK 145
QY   145 ATRVGINIFRLR 157

```



```

US-08/211-8333-2
Sequence 2, Application US/08211833
Patent No. 5489528
GENERAL INFORMATION:
APPLICANT: KOPETZKI, Erhard
APPLICANT: RUDOLPH, Rainer
APPLICANT: GROSSMANN, Adelbert
TITLE OF INVENTION: RECOMBINANT CORE STREPTAVIDIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelestein, Murray & Oram
STREET: 655 Fifteenth Street N.W., Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 2005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,833
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 41 35 543.1
FILING DATE: 28-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kitts, Monica C.
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-4013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

```

US-08-211-833-2

Query Match 16.0%; Score 135.5; DB 1; Length 128;
Best Local Similarity 32.8%; Pred. No. 1.6e-08;
Matches 42; Conservative 18; Mismatches 55; Indels 13; Gaps 7;
QY 36 KCSLTCKWTNDLCSNMTCGAVNSRCEFTCTVTAV-TATSNEIKESPLHCTONTINKRTO 94
DB 3 EAGITGTWYNQLGSTFIVTA-GADGALTGTYESAVGNAESRYVLGTGRYDSAPATDGSCT- 60
QY 95 PTEGFTVNWK----FSESTTVFTGQCFIDRNGKEV-LKTMWLLRSSVNDIGDDWKATRVG 149
DB 61 -ALGWTVAWKNNYRNAHSATTWSSGOV---CGAEARINTQWLLTSGTTE-ANAWKSTLWG 115
QY 150 INIFTRLR 157
DB 116 HDTFTKVK 123

RESULT 13

US-08-434-718-2
; Sequence 2, Application US/08434718
; Patent No. 5672691
; GENERAL INFORMATION:
; APPLICANT: KOPETZKI, Erhard
; APPLICANT: RUDOLPH, Rainer
; APPLICANT: GROSSMANN, Adelbert
; TITLE OF INVENTION: RECOMBINANT CORE STREPTAVIDIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikolaus, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W., Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,718
; FILING DATE: 4-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/211,833
; FILING DATE: 28-APRIL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 41 35 543.1
; FILING DATE: 28-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KILLS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P564-5007
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-434-718-2

Query Match 16.0%; Score 135.5; DB 1; Length 128;
Best Local Similarity 32.8%; Pred. No. 1.6e-08;
Matches 42; Conservative 18; Mismatches 55; Indels 13; Gaps 7;
QY 36 KCSLTCKWTNDLCSNMTCGAVNSRCEFTCTVTAV-TATSNEIKESPLHCTONTINKRTO 94

DB 3 EAGITGTWYNQLGSTFIVTA-GADGALTGTYESAVGNAESRYVLGTGRYDSAPATDGSCT- 60
QY 95 PTEGFTVNWK----FSESTTVFTGQCFIDRNGKEV-LKTMWLLRSSVNDIGDDWKATRVG 149
DB 61 -ALGWTVAWKNNYRNAHSATTWSSGOV---CGAEARINTQWLLTSGTTE-ANAWKSTLWG 115
QY 150 INIFTRLR 157
DB 116 HDTFTKVK 123

RESULT 14

US-08-831-399-16
; Sequence 16, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; TITLE OF INVENTION: Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-831-399-16

Query Match 16.0%; Score 135.5; DB 4; Length 128;
Best Local Similarity 32.8%; Pred. No. 1.6e-08;
Matches 42; Conservative 18; Mismatches 55; Indels 13; Gaps 7;

QY 36 KCSLTCKWTNDLCSNMTCGAVNSRCEFTCTVTAV-TATSNEIKESPLHCTONTINKRTO 94
DB 3 EAGITGTWYNQLGSTFIVTA-GADGALTGTYESAVGNAESRYVLGTGRYDSAPATDGSCT- 60
QY 95 PTEGFTVNWK----FSESTTVFTGQCFIDRNGKEV-LKTMWLLRSSVNDIGDDWKATRVG 149
DB 61 -ALGWTVAWKNNYRNAHSATTWSSGOV---CGAEARINTQWLLTSGTTE-ANAWKSTLWG 115
QY 150 INIFTRLR 157
; : : : :

Db 116 HDTFTKVK 123

RESULT 15
US-08-554-586-2
; Sequence 2, Application US/08554586
; Patent No. 5767379
; GENERAL INFORMATION:
; APPLICANT: BASZCZYNSKI, Chris
; APPLICANT: HOOD, Elizabeth
; APPLICANT: MADDOCK, Sheila
; APPLICANT: MEYER, Terry Euclaire
; APPLICANT: REGISTER III, James C.
; APPLICANT: WITCHER, Derrick
; APPLICANT: HOWARD, John A.
; TITLE OF INVENTION: COMMERCIAL PRODUCTION OF AVIDIN IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,586
; FILING DATE: 06-NOV-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/352/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-554-586-2

Query Match 13.6%; Score 115; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 ARKCSLTGKWTNDLGSNMTIG 54
DB 1 ARKCSLTGKWTNDLGSNMTIG 21

Search completed: September 17, 2002, 15:42:24
Job time: 165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:37:39 ; Search time 30.07 Seconds
(without alignments)
594.708 Million cell updates/sec

Title: US-09-743-690-7
Perfect score: 845
Sequence: 1 MESKFAHIIIVFLATPFET.....DWKATRVGINIFRLRTOKE 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	845	100.0	161	21 AAY44699	Potato proteinase
2	684	80.9	152	18 AAW29307	Wild-type avidin p
3	684	80.9	152	20 AAY17867	Chicken avidin. G
4	681	80.6	152	20 AAY30130	Amino acid sequenc
5	678	80.2	400	20 AAY03098	Bovine scavenger r
6	670	79.3	128	22 AAB35269	Chicken egg white
7	656	77.6	128	22 AAB35271	Chicken egg white
8	148	17.5	402	15 AAR56485	ScFV PRAS110 and p
9	145.5	17.2	183	19 AAW59217	S. avidinii strept
10	142.5	16.9	159	22 AAB35270	Streptomyces avidi
11	142.5	16.9	160	10 AAP93531	Mature streptavidin

12	142.5	16.9	161	21 AAY80515	Streptomyces avidi
13	142.5	16.9	162	21 AAY80514	Streptomyces avari
14	142.5	16.9	163	9 AAP80160	Biosynthetic prote
15	142.5	16.9	182	8 AAP70492	Streptavidin seque
16	142.5	16.9	183	7 AAP60625	Sequence of a stre
17	142.5	16.9	183	10 AAP93530	Streptavidin prote
18	142.5	16.9	183	14 AAR44491	Streptavidin gene.
19	142.5	16.9	183	18 AAW29306	Wild-type streptav
20	142.5	16.9	183	19 AAW59216	S. avidinii strept
21	142.5	16.9	183	20 AAY17868	Streptococcus stre
22	142.5	16.9	183	21 AAY44701	Streptavidin prote
23	142.5	16.9	183	22 AAB30692	Amino acid sequenc
24	142.5	16.9	184	21 AAY80513	Streptomyces avari
25	142.5	16.9	186	21 AAY80512	Streptomyces avari
26	142.5	16.9	412	22 AAB30694	A fusion of anti-C
27	142.5	16.9	415	15 AAR56484	ScFV PRAS109 and p
28	142.5	16.9	423	22 AAB30695	A fusion of anti-C
29	142.5	16.9	431	22 AAB30693	A fusion of single
30	142.5	16.9	435	15 AAR56483	ScFV PRAS108 and p
31	140.5	16.6	159	18 AAW29311	Streptavidin prote
32	139.5	16.5	188	21 AAY44700	Potato proteinase
33	139.5	16.5	183	19 AAW59218	S. avidinii strept
34	137.5	16.3	159	18 AAW29310	Streptavidin prote
35	137	16.2	186	21 AAY84021	Amino acid sequenc
36	136.5	16.2	159	18 AAW29309	Streptavidin prote
37	135.5	16.0	128	14 AAR34722	Core streptavidin.
38	135.5	16.0	128	17 AAW04211	Streptomyces avari
39	135.5	16.0	128	18 AAW29308	Recombinant Core-s
40	135.5	16.0	159	18 AAW29314	Streptavidin prote
41	135.5	16.0	270	20 AAY28928	S. avidini protein
42	135.5	16.0	673	17 AAW04208	Streptavidin/lucif
43	135	16.0	186	21 AAY84022	Amino acid sequenc
44	134.5	15.9	122	21 AAY80516	Streptomyces avari
45	134.5	15.9	122	21 AAY80517	Streptomyces avari

ALIGNMENTS

RESULT 1
AAY44699
ID AAY44699 standard; Protein: 161 AA.
XX AAY44699;
AC AAY44699;
XX 25-APR-2000 (first entry)
DT Potato proteinase inhibitor-I/Avidin fusion protein.
DE Potato proteinase inhibitor-I; pPI-I; streptavidin; worm; avidin;
XX Potato proteinase inhibitor-I; pPI-I; streptavidin; worm; avidin;
KW insect; plant-noxious protein; pest resistance; moth; insect; weevil;
KW grub; beetle; fly; thrip; locust; cricket; borer; mite; looper;
KW insecticidal; fusion protein.
XX Solanum tuberosum.
OS Synthetic.
XX Key Location/Qualifiers
FH Protein 1..31
FT /note= "Potato proteinase inhibitor-I sequence"
FT Peptide 1..23
FT Protein /note= "Potato proteinase inhibitor-I signal peptide"
34..161
/note= "Avidin sequence"
WO200004049-A1.
27-JAN-2000.
15-JUL-1999; 99WO-NZ00110.
15-JUL-1998; 98NZ-0331002.

DA (HORT-) HORTICULTURE & FOOD RES INST NEW ZEALAND.
 XX Christeller JT, Sutherland PW, Murray C, Markwick NP, Philip BA;
 PI Malone LA, Burgess EPJ;
 XX
 DR WPI: 2000-171244/15.
 DR N-PSDB: AA749865.
 XX
 PT New chimeric polypeptide and composition comprising the polypeptide
 PT useful for conferring pest resistance on plants -
 XX Example 2; Fig 8; 11pp; English.
 XX
 CC The present sequence is the potato proteinase inhibitor-I (PPI-I)/Avidin
 CC fusion protein. This comprises Avidin mature peptide, a plant-noxious
 CC protein, fused to the PPI-I signal peptide. Transformation of plant
 CC genome with the vector encoding the fusion protein can produce pest
 CC resistance in plants, plant derived products and stored harvest
 CC material. Pests that can be controlled include, cotton bollworm,
 CC tropical army-worm, European corn-borer or red mite, tobacco horn worm,
 CC loopers, rice stem borer, porina, cutworms, diamondback moth, potato
 CC tuber moth, codling moth, Indian meal moth, gypsy moth, argentine stem
 CC weevil, clover root weevil, grass-grubs, corn rootworm, rice and wheat
 CC weevils, mealworms, flour beetles, black field cricket, locusts,
 CC sawflies, Western flower thrips, Hessian flies or two-spotted mite.
 XX
 SQ Sequence 161 AA;

Query Match 100.0%; Score 845; DB 21; Length 161;
 Best Local Similarity 100.0%; Pred. No. 3.6e-83;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKEAHIIIVFLLATPEFELLARKESDGEIPARKSLTGKWTNDLGSNMTCGAVNSRG 60
 Db 1 meskfahiiivffllatpelfellarkesdgppeiparksltgkwtndlgsmctlgavnsrg 60
 QY 61 EFTGTITAVTATYSNEIKESPLHGHTQNTINKRQTPTFGFTVNMKFSSTVFTGQCQFIDR 120
 Db 61 efgtyitavtateneikesplhgtqntinkrqtptfgftvnmkfsstvtftgqcfidr 120
 QY 121 NGREVLKTMWMLRSSVNDIGDDWKATRVGINIFTRLRTQKE 161
 Db 121 ngrevlktmwmlrssvndigddwkatrvginiftrlrtqke 161

RESULT 2
 AAW29307
 ID AAW29307 standard; Protein: 152 AA.
 AC AAW29307;
 XX
 DT 27-APR-1998 (first entry)
 XX
 DE Wild-type avidin protein.
 XX
 KW Streptavidin; biotin; anti-interference reagent; detection; muten;
 XX avidin; non-specific binding.
 OS Unidentified.

XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal
 FT Protein 25..152
 XX
 PN DE19637718-A1.
 XX
 PD 02-OCT-1997.
 XX
 PF 16-SEP-1996; 96DE-1037718.
 XX
 PR 01-APR-1996; 96DE-1013053.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.
 XX Brandtetter H, Deger A, Engh R, Kopetzki E, Mueller R;
 PI Schmitt U;
 XX
 DR WPI: 1997-492043/45.
 DR N-PSDB: AA773194.
 XX
 PT Streptavidin and avidin muteins with reduced binding affinity for
 PT biotin - useful for reducing interference from nonspecific binding
 PT in assays
 XX
 PS Disclosure; Page 19-20; 26pp; German.
 XX
 CC This sequence represents an avidin which is used in a novel method
 CC of reducing interference from non-specific binding in assays. Muteins
 CC constructed from a core streptavidin or avidin sequence are selected that
 CC differ from the native polypeptide by at least one amino acid and have a
 CC binding affinity for biotin of less than 1010 1/mole. The biotin-bindable
 CC polypeptide may be present as a polymeric conjugate, e.g. with another
 CC polypeptide or protein, especially bovine serum albumin. These muteins
 CC are used as anti-interference reagents for reducing and/or avoiding
 CC nonspecific interactions in a process for detecting an analyte. In
 CC particular, they are used in assays where the streptavidin/avidin-biotin
 CC specific binding pair is involved for qualitative and/or quantitative
 CC determination of an analyte in a test sample, e.g. a heterogeneous
 CC immunoassay or a hybridisation assay. Despite having a lower binding
 CC affinity for biotin, the muteins have high immunological cross-reactivity
 CC with native streptavidin and avidin.
 XX
 SQ Sequence 152 AA;

Query Match 80.9%; Score 684; DB 18; Length 152;
 Best Local Similarity 97.7%; Pred. No. 7.6e-66;
 Matches 129; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 30 PEIPARKSLTGKWTNDLGSNMTCGAVNSRGFTGTITAVTATYSNEIKESPLHGHTQNTI 89
 Db 21 pglisarksltgkwtndlgsmctlgavnsrgftgtyitavtatsneikesplhgtqnti 80
 QY 90 NKRTQPTFGFTVNMKFSSTVFTGQCQFIDRNGKEVLKTMWMLRSSVNDIGDDWKATRVG 149
 Db 81 nkrtqptfgftvnmkfsstvtftgqcfidrngkevlktmwmlrssvndigddwkatrvg 140
 QY 150 INIFTRLRTQKE 161
 Db 141 iniftrlrtqke 152

RESULT 3
 AAY17867
 ID AAY17867 standard; Protein: 152 AA.
 AC AAY17867;
 XX
 DT 20-AUG-1999 (first entry)
 XX
 DE Chicken avidin.
 XX
 KW Avidin; streptavidin; batroxobin; fibrinogen converting enzyme;
 XX hybrid; fusion protein; sealant; surgery; reduce bleeding; fibrin.
 XX
 OS Gallus gallus.
 XX
 PN WO9929838-A1.
 XX
 PD 17-JUN-1999.
 XX
 PF 09-DEC-1998; 98WO-US26086.
 XX
 PR 09-DEC-1997; 97US-0067978.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX Cederholm-Williams SA;
 PI WPI: 1999-385599/32.
 DR N-PSDB; AAX80197.
 XX A fibrinogen-converting enzyme fusion protein
 PT Disclosure; Page 27; 35pp; English.
 PS
 XX The present invention describes a fibrinogen-converting enzyme fusion
 CC protein (FCE). The fusion protein is a multidomain protein comprising:
 CC (a) a FCE; and (b) a first member of a binding pair, that is linked to
 CC the FCE chain; (i) directly by bonds utilizing the N-terminal amino
 CC groups, the C-terminal carboxy groups or side-chain functionalities;
 CC (ii) via a bifunctional linkage moiety linking the groups or
 CC functionalities; or (iii) by the first member binding to the second
 CC member of the binding pair, where the second member of the binding pair
 CC is covalently attached to the first polypeptide chain. The FCE can be
 CC used in a method for producing fibrin. Fibrin is useful as a sealant in
 CC surgery to, e.g. reduce bleeding by sealing blood vessels, and tissues
 CC that have been dissected either in surgery or through wounding. The
 CC fusion protein allows for the removal of the fibrinogen converting
 CC enzyme from the fibrin sealant preparation via the binding of
 CC streptavidin to a biotin solid support. The present sequence represents
 CC chicken (Gallus gallus) avidin as given in the present invention.
 XX
 SQ Sequence 152 AA;
 Query Match 80.9%; Score 684; DB 20; Length 152;
 Best Local Similarity 97.7%; Pred. No. 7.6e-66;
 Matches 129; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 30 PEIPARKSLTGKWTNDLGSNMFTIGAVNSRGFTGTYYTAVTATSNKEIKESPLHGTONTI 89
 Db 21 pglarkscsltgkwtndlgsnmftigavnsrgftgttytattatsneikesplhgtonti 80
 QY 90 NKRTQPTFGFTVNMKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
 Db 81 nkrtqptfgftvnmkfsstvtftgqcfidrngkevlktmwlrrssvndigddwkatrvg 140
 QY 150 INIFTRLTQKE 161
 Db 141 iniftrlrtqke 152
 RESULT 4
 ID AAY30130 standard; peptide; 152 AA.
 AC AAY30130;
 DT 26-OCT-1999 (first entry)
 XX Amino acid sequence of the avidin monomer.
 DE Avidin; biotin; neutravidin; ligand-binding molecule;
 KW cross-linked structure; polynucleotide delivery; DNA linking;
 KW biotin-avidin networked gene system; BANG system; DNA vaccine.
 XX Unidentified.
 OS
 FH Key Location/Qualifiers
 FX Peptide 1..24
 FT Disulfide-bond 28..107
 FT Modified-site 41
 FT /note= "linked to a carbohydrate chain"
 FT Misc-difference 58
 FT /note= "this residue may also be Thr"

FT Misc-difference 57
 FT /note= "biotin binding occurs here"
 XX WO9939744-A1.
 XX 12-AUG-1999.
 PD 10-FEB-1999; 99WO-US02673.
 XX 10-FEB-1998; 98US-0074213.
 XX (OHIS) UNIV OHIO STATE RES FOUND.
 PA Luo D, Muller MT;
 PI WPI: 1999-518369/43.
 DR New cross-linked polynucleotide complexes, useful for cell-targeted
 XX polynucleotide delivery
 PT Example 1; Page 40; 135pp; English.
 PS
 XX The present sequence represents an avidin monomer. Avidin exists as a
 CC tetramer, where each of the four monomer units binds to one biotin.
 CC Biotin is complexed with a neutral avidin (neutravidin), and then
 CC covalently coupled with polynucleotides, to exemplify the invention. The
 CC specification describes a composition of complexes of polynucleotide
 CC molecules covalently coupled to ligand moieties that are specifically
 CC bound to ligand-binding molecules to form a cross-linked structure.
 CC The composition allows incorporation of diverse oligonucleotides or
 CC polynucleotides into a single complex for concomitant delivery into the
 CC same cell. The composition is used to deliver polynucleotides to viable
 CC cells. The method and compositions provide a new way of linking DNA
 CC molecules and are useful for gene over-expression and non-covalent
 CC cloning. The new system is called biotin-avidin networked gene (BANG)
 CC system. It is possible that the BANG system can be used to link multiple
 CC gene complexes to elicit broader immune reactions, e.g. as DNA vaccines.
 CC The BANG system can also be used as a cloning tool. The BANG system also
 CC eliminates the reading frame shift and cloning problems associated with
 CC plasmids and vectors.
 XX
 SQ Sequence 152 AA;
 Query Match 80.6%; Score 681; DB 20; Length 152;
 Best Local Similarity 97.0%; Pred. No. 1.6e-65;
 Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 30 PEIPARKSLTGKWTNDLGSNMFTIGAVNSRGFTGTYYTAVTATSNKEIKESPLHGTONTI 89
 Db 21 pglarkscsltgkwtndlgsnmftigavnsrgftgttytattatsneikesplhgtonci 80
 QY 90 NKRTQPTFGFTVNMKFSSTVFTGQCIFDRNGKEVLKTMWLLRSSVNDIGDDWKATRVG 149
 Db 81 nkrtqptfgftvnmkfsstvtftgqcfidrngkevlktmwlrrssvndigddwkatrvg 140
 QY 150 INIFTRLTQKE 161
 Db 141 iniftrlrtqke 152
 RESULT 5
 ID AAY03098 standard; Protein; 400 AA.
 AC AAY03098;
 XX
 XX 03-DEC-1999 (first entry)
 DT Bovine scavenger receptor class A (ScR)/avidin fusion protein.
 XX Scavenger receptor class A; ScR; avidin; fusion protein; bovine; ECD;
 KW membrane-spanning domain; extracellular domain; biotin-binding activity;
 KW

KW endocytosis.
 XX Synthetic.
 OS BOS taurus.
 XX
 PN WO9942577-A2.
 XX
 DD 26-AUG-1999.
 XX
 XX 23-FEB-1999; 99WO-GB00546.
 XX
 XX 23-FEB-1998; 98GB-0003757.
 PR 24-JUN-1998; 98GB-0013653.
 XX
 XX (EURO-) EUROGENE LTD.
 PA
 XX
 XX
 PI Vira-Harttuala S, Kulomaa M, Lehtolainen P, Marjomaki V, Airenne K;
 XX
 DR WPI; 1999-561345/47.
 DR N-PSDB; AAZ09997.
 XX
 XX
 XX
 PT NEW fusion proteins having an extracellular domain with biotin-binding
 PT activity, used to target biotinylated molecules to specific sites in
 PT tissues -
 XX
 XX
 PS Claim 5; Page 21-23; 23pp; English.
 XX
 CC This invention describes a novel protein (A) which comprises a
 CC membrane-spanning domain and an extracellular domain (ECD), where the
 CC ECD comprises biotin-binding activity. Using the proteins or encoding
 CC nucleic acid molecules it is possible to target biotinylated molecules
 CC to specific sites in tissues. Molecules targeted in this way may be
 CC taken up by the tissues or cells by endocytosis, allowing the molecules
 CC to exert their effects within or on the cell. This sequence represents a
 CC bovine scavenger receptor class A/avidin fusion protein which is used in
 CC the description of the invention.
 XX
 XX
 SO Sequence 400 AA;
 }
 Query Match 80.2%; Score 678; DB 20; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.2e-64;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 ARKCSLCKWPNLGLSNMTIGAVNSRCEFTGTCTVAVTATSNKESPLHGTONTINKRT 93
 DD 273 ARKCSLTKWLNGLSNMTIGAVNSRGFTGTCTVAVTATSNKESPLHGTONTINKRT 332
 QY 94 QPTFGFTVNMKFSSTVFTGQCFTDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIF 153
 Db 333 qptfgftvnmkfsstvtftgqcfidrngkevlktmwlrrssvndigddwkatrvginif 392
 QY 154 TRLRTOKE 161
 Db 393 trlrrtqke 400
 RESULT 6
 AAB35269
 ID AAB35269 standard; protein; 128 AA.
 XX
 AC AAB35269;
 XX
 DT 08-MAY-2001 (first entry)
 XX
 DE Chicken egg white avidin.
 XX
 KW Chicken; egg white; avidin; bacterial streptavidin; mutant; mutein;
 KW dimer; biotin binding activity; protein purification; immunoassay;
 KW signal transduction.
 XX
 OS Gallus gallus.
 OS Synthetic.
 XX
 PN WO200105977-A1.
 XX
 PD 25-JAN-2001.
 XX
 PF 13-JUL-2000; 2000WO-US18959.
 XX
 PR 15-JUL-1999; 99US-0354097.
 XX
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA
 PI Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;
 DR WPI; 2001-147344/15.
 XX
 XX Avidin and streptavidin mutants comprising tryptophan 110 and
 PT tryptophan 120 respectively, substituted by lysine for use in all known
 PT avidin/streptavidin-biotin systems and a purification process -
 XX
 XX Disclosure; Page 43; 49pp; English.
 XX
 CC The present invention provides a mutant chicken egg white avidin protein
 CC and a mutant bacterial streptavidin, which are capable of generating
 CC stable dimers with reversible biotin binding activity. These can be used
 CC in signal transduction studies, protein purification and immunoassays.
 CC The present sequence is the chicken egg white avidin protein.
 XX
 XX
 SO Sequence 128 AA;
 Query Match 79.3%; Score 670; DB 22; Length 128;
 Best Local Similarity 98.4%; Pred. No. 2e-64;
 Matches 126; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 34 ARKCSLTGKWTNDLGSNMTIGAVNSRGFTGTCTVAVTATSNKESPLHGTONTINKRT 93
 Db 1 arkcsltgkwtndlgsnmtigavnsrgftgtctvtatatsnikesplhgtontinkrt 60
 QY 94 QPTFGFTVNMKFSSTVFTGQCFTDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIF 153
 Db 61 qptfgftvnmkfsstvtftgqcfidrngkevlktmwlrrssvndigddwkatrvginif 120
 QY 154 TRLRTOKE 161
 Db 121 trlrrtqke 128
 RESULT 7
 AAB35271
 ID AAB35271 standard; protein; 128 AA.
 XX
 AC AAB35271;
 XX
 DT 08-MAY-2001 (first entry)
 XX
 DE Chicken egg white avidin W10K mutant.
 XX
 KW Chicken; egg white; avidin; bacterial streptavidin; mutant; mutein;
 KW dimer; biotin binding activity; protein purification; immunoassay;
 KW signal transduction.
 XX
 OS Gallus gallus.
 OS Synthetic.
 XX
 PN WO200105977-A1.
 XX
 PD 25-JAN-2001.
 XX
 PF 13-JUL-2000; 2000WO-US18959.
 XX
 PR 15-JUL-1999; 99US-0354097.
 XX
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA
 PI Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;
 DR WPI; 2001-147344/15.
 XX
 XX Avidin and streptavidin mutants comprising tryptophan 110 and
 PT tryptophan 120 respectively, substituted by lysine for use in all known
 PT avidin/streptavidin-biotin systems and a purification process -
 XX
 XX Disclosure; Page 43; 49pp; English.
 XX
 CC The present invention provides a mutant chicken egg white avidin protein
 CC and a mutant bacterial streptavidin, which are capable of generating
 CC stable dimers with reversible biotin binding activity. These can be used
 CC in signal transduction studies, protein purification and immunoassays.
 CC The present sequence is the chicken egg white avidin protein.
 XX
 XX
 SO Sequence 128 AA;

PN WO200105977-A1.
 XX
 PD 25-JAN-2001.
 XX
 PF 13-JUL-2000; 2000WO-US18959.
 XX
 PR 15-JUL-1999; 99US-0354097.
 XX
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA
 PI Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;
 DR WPI; 2001-147344/15.
 XX
 XX Avidin and streptavidin mutants comprising tryptophan 110 and
 PT tryptophan 120 respectively, substituted by lysine for use in all known
 PT avidin/streptavidin-biotin systems and a purification process -
 XX
 XX Disclosure; Page 43; 49pp; English.
 XX
 CC The present invention provides a mutant chicken egg white avidin protein
 CC and a mutant bacterial streptavidin, which are capable of generating
 CC stable dimers with reversible biotin binding activity. These can be used
 CC in signal transduction studies, protein purification and immunoassays.
 CC The present sequence is the chicken egg white avidin protein.
 XX
 XX
 SO Sequence 128 AA;
 Query Match 79.3%; Score 670; DB 22; Length 128;
 Best Local Similarity 98.4%; Pred. No. 2e-64;
 Matches 126; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 34 ARKCSLTGKWTNDLGSNMTIGAVNSRGFTGTCTVAVTATSNKESPLHGTONTINKRT 93
 Db 1 arkcsltgkwtndlgsnmtigavnsrgftgtctvtatatsnikesplhgtontinkrt 60
 QY 94 QPTFGFTVNMKFSSTVFTGQCFTDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIF 153
 Db 61 qptfgftvnmkfsstvtftgqcfidrngkevlktmwlrrssvndigddwkatrvginif 120
 QY 154 TRLRTOKE 161
 Db 121 trlrrtqke 128
 RESULT 7
 AAB35271
 ID AAB35271 standard; protein; 128 AA.
 XX
 AC AAB35271;
 XX
 DT 08-MAY-2001 (first entry)
 XX
 DE Chicken egg white avidin W10K mutant.
 XX
 KW Chicken; egg white; avidin; bacterial streptavidin; mutant; mutein;
 KW dimer; biotin binding activity; protein purification; immunoassay;
 KW signal transduction.
 XX
 OS Gallus gallus.
 OS Synthetic.
 XX
 PN WO200105977-A1.
 XX
 PD 25-JAN-2001.
 XX
 PF 13-JUL-2000; 2000WO-US18959.
 XX
 PR 15-JUL-1999; 99US-0354097.
 XX
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA
 PI Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;
 DR WPI; 2001-147344/15.
 XX
 XX Avidin and streptavidin mutants comprising tryptophan 110 and
 PT tryptophan 120 respectively, substituted by lysine for use in all known
 PT avidin/streptavidin-biotin systems and a purification process -
 XX
 XX Disclosure; Page 43; 49pp; English.
 XX
 CC The present invention provides a mutant chicken egg white avidin protein
 CC and a mutant bacterial streptavidin, which are capable of generating
 CC stable dimers with reversible biotin binding activity. These can be used
 CC in signal transduction studies, protein purification and immunoassays.
 CC The present sequence is the chicken egg white avidin protein.
 XX
 XX
 SO Sequence 128 AA;

PI Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;

DR WPI; 2001-147344/15.

PT Avidin and streptavidin mutants comprising tryptophan 110 and
PT tryptophan 120 respectively, substituted by lysine for use in all known
PT avidin/streptavidin-biotin systems and a purification process -

XX Disclosure; Page 45; 49pp; English.

XX The present invention provides a mutant chicken egg white avidin protein
CC and a mutant bacterial streptavidin, which are capable of generating
CC stable dimers with reversible biotin binding activity. These can be used
CC in signal transduction studies, protein purification and immunoassays.
CC The present sequence is the mutant chicken egg white avidin protein.

XX Sequence 128 AA;

Query Match 77.6%; Score 656; DB 22; Length 128;
Best Local Similarity 97.7%; Pred. No. 6.3e-63;
Matches 125; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 34 ARKSLTGKWTNDLGSNMTIGAVNSRGFTGYITAVTATSNEIKESPLHGQTONTINKRT 93
Db 1 arkcsltgkwtndlgsnmtigavnsrgftgyttavtatsneikesplhgtentinkrt 60
Qy 94 QPTGFTVNWKSESTVFTGQCFTDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIF 153
Db 61 qptgftvnwkfsestvtftgqcfidrngkevlktmwlrrssvndigddkkatrvginif 120
Qy 154 TRLRTOKE 161
Db 121 trlrltqke 128

RESULT 8

AAR56485
ID AAR56485 standard; Protein; 402 AA.

XX AC AAR56485;

XX DT 26-MAR-1995 (first entry)

XX DE ScFv PRAS110 and PRAS114.

XX KW Amplification; single chain variable region fusion protein; PCR.

XX OS Synthetic.

XX PN WO9415644-A.

XX PD 21-JUL-1994.

XX PF 17-JAN-1994; 94WO-GB00087.

XX PR 15-JAN-1993; 93GB-0000686.

XX PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

XX PI Deonara M, Epenetos AA, Spooner RA;

XX DR WPI; 1994-248907/30.

DR N-PSDB; AAQ70661.

XX New cpgs, comprising a targeting portion and a cytotoxic portion
PT - used esp. for treating mammals for destroying target cells,
PT partic. tumour cells

XX Disclosure; Fig 11; 114pp; English.

XX The sequence is that of the ScFv PRAS110 and PRAS114 between HindIII and
CC EcoRI site obtd. by PCR.

CC See also AAR56482-4.

XX Sequence 402 AA;

Query Match 17.5%; Score 148; DB 15; Length 402;
Best Local Similarity 34.1%; Pred. No. 1.8e-07;
Matches 47; Conservative 18; Mismatches 57; Indels 16; Gaps 8;

Qy 29 GPEIPA---RKCSLTGKWTNDLGSNMTIGAVNSRGFTGYITAV-TATSNEIKESPLHG 84
Db 267 gleapaapaegaigtwynglsgflvta-gadgaltgyesavgnaesryvitgryds 325
Qy 85 TQNTINKRTQPTGFTVNWK----FSBSTVFTGQCFTDRNGKEV-LKTMWLLRSSVNDI 139
Db 326 apatdgsrt--algwtvawknnrynahsattsggyv---ggaeaintqltsgtte- 379
Qy 140 GDDWKATRVGINIFRLR 157
Db 380 anawkstlvghdtftkvvk 397

RESULT 9

AAW59217

ID AAW59217 standard; Protein; 183 AA.

XX AC AAW59217;

XX DT 27-AUG-1998 (first entry)

XX DE S. avidinii streptavidin mutant protein #1.

XX KW Streptavidin; ligand; binding affinity; mutant; isolation;

XX KW purification; recover; immobilise.

XX OS Streptomyces avidinii.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 68..71

FT /label= VTAR68ESAV

FT /note= "Wild type ESAV is replaced by VTAR. Numbering
is from the start of the mature protein"

XX PN EP835934-A2.

XX PD 15-APR-1998.

XX PF 09-OCT-1997; 97EP-0117504.

XX PR 10-OCT-1996; 96DE-1041876.

XX PA (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.

XX PI Skerra A, Voss S;

XX DR WPI; 1998-218868/20.

DR N-PSDB; AAV34715.

XX Streptavidin mutants with higher binding affinity for peptide

PT ligands - have mutation in amino acid region 44-53, used to isolate,

PT purify or determine fusion proteins including these ligands

XX Disclosure; Page -; 21pp; German.

XX This sequence represents a mutant streptavidin protein isolated from
CC Streptomyces avidinii where the residues ESAV at position 44-47 of
CC the mature wild type sequence are replaced by VTAR. This sequence is
CC used to produce mutants which are used in a method to assay the binding
CC affinity of streptavidin mutants. These mutants have a mutation within
CC the amino acid (aa) region 44-53 of the wild-type protein show a higher
CC binding affinity than the wild-type for peptide ligands that include
CC the sequence of formula Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa, Y

CC and Z are both Gly, or Y = Glu and Z = Arg or Lys. Recombinant streptavidin mutants can be used to isolate, purify and determine proteins or to determine/recover substances that contain streptavidin-binding groups. Such compounds may also be used to immobilise fusions on microtitre plates, microbeads or sensor chips. CC NOTE: This sequence does not appear in the specification but has been constructed from the wild-type streptavidin sequence represented in AAV34714.

XX Sequence 183 AA;

Query Match 17.2%; Score 145.5; DB 19; Length 183;
Best Local Similarity 32.3%; Pred. NO. 1.1e-07;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

QY 31 EIPARKSLTCKWTNDLGSNMFTICAVNSRGFTGTITAV-TATSNKESPLHGTQNTI 89
DB 33 QVSADEAGILGTWYNQIGSTFLVTA-GADGALTGYTARGNAESRYVLTGRYDAPATD 91
QY 90 NKRTQPTFGFTVNNK----FSESTVFTGQCFTIDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
DB 92 GSGLT--ALGWTVAWKNNYRNHSAATTWSGGYV---GGAEARINTQWLLTSGLTTE-ANAWK 145
QY 145 ATRVGINIFTRLR 157
DB 146 STLVGHDTFTKVK 158

RESULT 10
AAB35270
ID AAB35270 standard; protein: 159 AA.

XX AAB35270;
AC
XX 08-MAY-2001 (first entry)

XX Streptomyces avidinii streptavidin.

XX Chicken; egg white; avidin; bacterial streptavidin; mutant; mutein;
KW dimer; biotin binding activity; protein purification; immunoassay;
KW signal transduction.

XX Streptomyces avidinii.

XX WO200105977-A1.

XX 25-JAN-2001.

XX 13-JUL-2000; 2000WO-US18959.

XX 15-JUL-1999; 99US-0354097.

XX (VEDA) VEDA RES & DEV CO LTD.

XX Kulomaa MS, Bayer EA, Wilchek M, Laitinen OH;

XX WPI; 2001-147344/15.

XX Avidin and streptavidin mutants comprising tryptophan 110 and PT tryptophan 120 respectively, substituted by lysine for use in all known PT avidin/streptavidin-biotin systems and a purification process -

XX Disclosure; Page 44; 49pp; English.

XX The present invention provides a mutant chicken egg white avidin protein CC and a mutant bacterial streptavidin, which are capable of generating CC stable dimers with reversible biotin binding activity. These can be used CC in signal transduction studies, protein purification and immunoassays. CC The present sequence is the Streptomyces avidinii streptavidin protein.

XX Sequence 159 AA;

Query Match 16.9%; Score 142.5; DB 22; Length 159;
Best Local Similarity 32.3%; Pred. NO. 2e-07;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

QY 31 EIPARKSLTCKWTNDLGSNMFTICAVNSRGFTGTITAV-TATSNKESPLHGTQNTI 89
DB 9 QVSADEAGILGTWYNQIGSTFLVTA-GADGALTGYTARGNAESRYVLTGRYDAPATD 67
QY 90 NKRTQPTFGFTVNNK----FSESTVFTGQCFTIDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
DB 68 GSGLT--ALGWTVAWKNNYRNHSAATTWSGGYV---GGAEARINTQWLLTSGLTTE-ANAWK 121
QY 145 ATRVGINIFTRLR 157
DB 122 STLVGHDTFTKVK 134

RESULT 11

AAP93531

ID AAP93531 standard; protein: 160 AA.

XX AAP93531;

XX 04-JUN-1990 (first entry)

XX Mature streptavidin as encoded by a synthetic gene.

XX Streptavidin; synthetic gene; Streptomyces avidinii;

XX Biotin; chimeric gene.

XX WO8903422-A.

XX 20-APR-1989.

XX 07-OCT-1988; 88WO-GB00831.

XX 08-OCT-1987; 87GB-0023661.

XX (BRBI-) BRIT BIO-TECHN LTD.

XX Edwards RW;

XX WPI; 1989-130040/17.

XX N-PSDB; AAN90755.

XX DNA sequence encoding streptavidin and vector -

XX comprising hybrid gene encoding fusion protein with

XX biotin-binding activity

XX Fig 2; page 2/5; 22pp; English.

XX The synthetic gene is constructed from 24 oligomers (BB214-BB237)

XX which were synthesised by automated phosphoramidite chemistry.

XX Apart from Met-1, it has the same sequence as natural mature streptavidin

XX protein produced by Streptomyces avidinii.

XX Sequence 160 AA;

Query Match 16.9%; Score 142.5; DB 10; Length 160;
Best Local Similarity 32.3%; Pred. NO. 2e-07;
Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

QY 31 EIPARKSLTCKWTNDLGSNMFTICAVNSRGFTGTITAV-TATSNKESPLHGTQNTI 89
DB 10 QVSADEAGILGTWYNQIGSTFLVTA-GADGALTGYTARGNAESRYVLTGRYDAPATD 68
QY 90 NKRTQPTFGFTVNNK----FSESTVFTGQCFTIDRNGKEV-LKTMWLLRSSVNDIGDDWK 144
DB 69 GSGLT--ALGWTVAWKNNYRNHSAATTWSGGYV---GGAEARINTQWLLTSGLTTE-ANAWK 122
QY 145 ATRVGINIFTRLR 157

Db 124 stlvahdtftkvk 136

RESULT 13
AAY80514
ID AAY80514 standard; Protein: 162 AA.

...
AC AAY80514;
XX
DT 06-JUN-2000 (first entry)

DT 06-JUN-2000 (first entry)

Streptomyces avidinii prost protein.

Plant somatic tissue degeneration; plant essential factor; depletion;
viability; prost gene; plant development; plant morphology; flower;
fruit plant.

OS Streptomyces avidinii.

XX
PN
WQ200007427-A2

XX 17 22 23 24 25 26 27 28 29 30

XX

PF 30-JUL-1999; 93WO-IL00420: XX

PR 03-AUG-1998; 98IL-0125632.

PA (AGRI-) AGRIC RES ORG.

PI Kapulnik Y, Ginzberg I;

XX
DP
WPT: 2000-195402/17

DR N-PSDB; AAZ91075.

PT Degeneration of somatic plant tissue by expression of a heterologous
PT protein, useful for controlling plant development and morphology, such
PT as decreasing the number of flowers present to increase the number of
PT fruit -

Examples: Page 86: 91pp: English.

The invention relates to a method of effecting degeneration of a somatic plant tissue by expressing a heterologous protein capable of binding a plant essential factor (PEF), in somatic plant tissue cells, where heterologous protein expression causes depletion of the PEF so the plant viability is maintained, while simultaneous degeneration of the somatic plant tissue is effected. This sequence represents the Streptomyces avidinii prst protein as an example of a heterologous protein introduced into the plants. The methods can provide for the selective and optionally reversible cell degeneration in somatic plant tissue. They can be used for artificially controlling plant development and morphology. They can be used e.g. to decrease the number of flowers in fruit producing plants so as to increase the number of fruits which reach maturity.

Sequence	162 AA;
SQ	

Query Match 16.9%; Score 142.5; DB 21; Length 162;
Best Local Similarity 32.3%; Pred. No. 2e-07;

Matches 43; Conservative 20; Mismatches 57; Indels 13; Gaps 7;

Qy	31	ETPARKCSLTGKWINDLGSNNITIGAVNSRGEFTGYITAV-TATSNIEIKESPLHGQTNTI	89
Db	12	qvsaaeagitgtwnqlgstfivta-gadgalgtgyesavgnaesryvltgrydsapad	70

Qy 90 NKRTPTFGTWNK----FSESTVFTGQCFFIDRNGKEV-LKTMWLLRSSVNDIGDDWK 144

Db 71 gsqt--algwtvawknynrnahsattswgqyv---ggaeairintqwltsgtte-anawk 124

Oy 145 ATRVGINIFTRLR 157

DB 125 stlva bdt fr kvk 137

